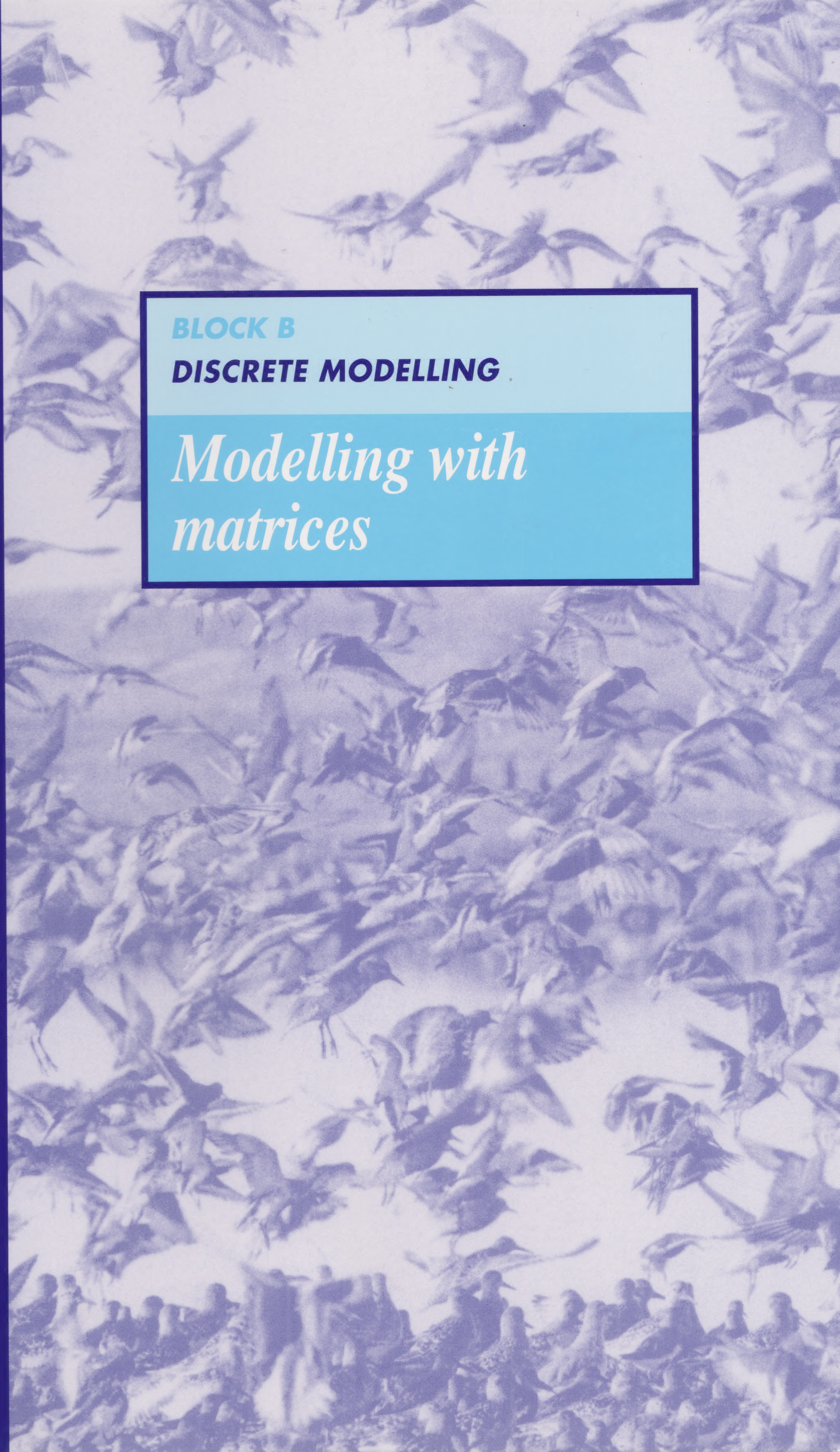
MST121 Chapter B2

The Open University

A first level interdisciplinary course

Using **Mathematics**



BLOCK B

DISCRETE MODELLING

Modelling with matrices

CHAPTER

B2

MST121 Chapter B2

The Open University

A first level Interdisciplinary course

Using Mathematics

BLOCK B

DISCRETE MODELLING

Modelling with matrices

Prepared by the course team

CHAPTER

B2

About this course

This course, MST121 *Using Mathematics*, and the courses MU120 *Open Mathematics* and MS221 *Exploring Mathematics* provide a flexible means of entry to university-level mathematics. Further details may be obtained from the address below.

MST121 uses the software program Mathcad (MathSoft, Inc.) and other software to investigate mathematical and statistical concepts and as a tool in problem solving. This software is provided as part of the course.

This publication forms part of an Open University course. Details of this and other Open University courses can be obtained from the Course Information and Advice Centre, PO Box 724, The Open University, Milton Keynes, MK7 6ZS, United Kingdom: tel. +44 (0)1908 653231, e-mail ces-gen@open.ac.uk

Alternatively, you may visit the Open University website at http://www.open.ac.uk where you can learn more about the wide range of courses and packs offered at all levels by The Open University.

To purchase this publication or other components of Open University courses, contact Open University Worldwide Ltd, The Open University, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom: tel. +44 (0)1908 858785, fax +44 (0)1908 858787, e-mail ouwenq@open.ac.uk, website http://www.ouw.co.uk

The Open University, Walton Hall, Milton Keynes MK7 6AA.

First published 1997. Second edition 2002.

Edited, designed and typeset by The Open University, using the Open University TEX System.

Printed and bound in the United Kingdom by The Burlington Press, Foxton, Cambridge CB2 6SW.

ISBN 0 7492 3464 4

2.1

Contents

Study guide

There are five sections in this chapter. They are intended to be studied consecutively in five study sessions.

Section 1, which starts and ends with video sequences, requires the use of a video player, and Section 4 requires the use of the computer together with Computer Book B.

The pattern of study for each session might be as follows.

Study session 1: Section 1.

Study session 2: Section 2.

Study session 3: Section 3.

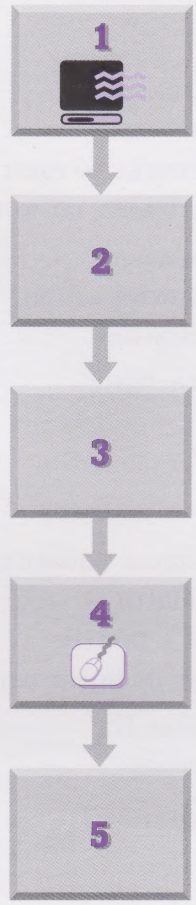
Study session 4: Section 4.

Study session 5: Section 5.

Each of the sections should take two to three hours of study.

Section 1 introduces matrices, and Section 2 covers a little of their theory. Sections 3 and 4 contain a modelling case-study that shows one application of matrices, whilst Section 5 continues with more of the theory of matrices. An alternative study pattern, which you might prefer, is to study Section 5 before Section 3, thus putting all the theory together before studying an application.

In addition to the work outlined above, there is an *optional* band of Audio Tape 2, in which various applications of matrices are described. The best time to listen to this band is after you have completed your study of the chapter.



Introduction

As its title suggests, this chapter introduces and uses mathematical objects called *matrices*. Matrices are a very important topic in mathematics and find widespread use in mathematical modelling. In fact, almost every problem that is too complicated to be solved with paper and pencil is first reduced to a matrix problem before being solved numerically by computer.

The singular of *matrices* is *matrix*.

Matrices are introduced in Section 1 through the concept of *networks*. The network image provides a way of visualising matrices and how to multiply them. In Section 2, the matrix concept is built upon and standard matrix notation is defined.

In Section 3, a population problem is studied. This problem can be solved using matrices with the basic operations defined in Section 2. The model which we shall use is a sequence model similar to one of those used in the previous chapter, but with the population split into subpopulations.

Section 4 is a computer session designed to show how useful computer software is for dealing with matrices. The section starts by introducing the ways in which the computer can manipulate matrices, then proceeds to investigate further the population problem from Section 3.

In Section 5, pairs of simultaneous linear equations are expressed in terms of matrices and solved, where possible, by manipulating those matrices.

1 Networks and matrices

To study this section you will need a video player and the Video Tape. The section starts with a video sequence (Subsection 1.1) and concludes with one (Subsection 1.3).

In this section you will meet mathematical objects called *matrices*, and see how to 'multiply' two matrices. Matrices can be used to represent particular kinds of *network*, and 'combining' two networks corresponds to *matrix multiplication*.

1.1 Networks

This section starts with a discussion of networks, on the Video Tape.

Watch Video Band B, 'Networks and matrices', until you reach the caption 'Stop tape 1'.

Two of the main ideas from the video are the following.

◇ A **network diagram** is a mathematical representation of a physical network. Each point at which a network diagram branches is called a **node**. Two nodes in a network can be connected by a **pipe**.

◇ In the networks that we consider, the pipes are labelled by numbers. Suppose that 1 litre of water is input at a node, and that it flows from the node down a number of pipes. The number labelling each pipe is the amount (in litres) of that 1 litre of water which emerges via that pipe at the node at the bottom of the pipe.

Where no confusion arises, the word 'network' is used to mean either 'network diagram' or 'physical network'. Similarly, the word 'pipe' will mean either 'pipe in a network diagram' or 'physical pipe'.

The first network introduced in the video can be represented by the following network diagram.

Note that two pipes that cross on a diagram do not join with each other unless the crossing point is marked as a node, by means of a circle. For example, the pipe from node A to node V does not join with the pipe from node B to node U, because the crossing point is not marked as a node.

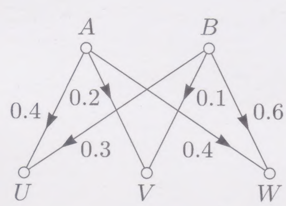


Figure 1.1 Network with two input nodes (A and B), three output nodes (U, V and W) and six pipes

Consider what happens when 1 litre of water is input at the node labelled A. The water flows downwards, and can be collected at the nodes labelled U, V and W. The amount of water reaching each node can be read off the network diagram by using the pipe labels. So in this case 0.4 litres of water reach node U, 0.2 litres of water reach node V, and 0.4 litres of water reach node W.

Now try the following activity.

Activity 1.1 Calculating network outputs

Consider the network diagram shown in Figure 1.1.

(a) How much water is output at the nodes U, V and W if the only input is 1 litre of water at node B?

(b) How much water is output at the nodes U, V and W if the only input is 2 litres of water at node A?

Solutions are given on page 46.

In Activity 1.1(b), you should have found that doubling the amount of water input doubles the amount of water at each output node. The same principle applies if water is input at both node A and node B: the amount of water in an output node is the sum of the water reaching that node from node A and from node B. Consider, for example, what happens when 1 litre of water is input at node A and another 1 litre is input at node B. At node U, 0.4 litres of water arrive from node A and 0.3 litres of water arrive from node B. So, in total, there will be 0.7 litres of water in the container at node U. Similarly, $0.2 + 0.1 = 0.3$ litres of water collect at node V, and $0.4 + 0.6 = 1$ litre of water collects at node W.

Activity 1.1(b) indicates that the label on the pipe from node A to node U (0.4), for example, gives the *proportion* of water input at node A that reaches node U. The other pipe labels can be interpreted in the same way.

Now try using this reasoning yourself to calculate network outputs.

Activity 1.2 Calculating network outputs

Consider the network diagram shown in Figure 1.1.

(a) How much water is output at the nodes U, V and W if 1 litre of water is input at node A and 2 litres of water are input at node B?

(b) How much water is output at the nodes U, V and W if x litres of water are input at node A and y litres of water are input at node B?

Solutions are given on page 46.

Activity 1.3 Calculating more network outputs

Consider the following network diagram, which represents the second network shown in the video.

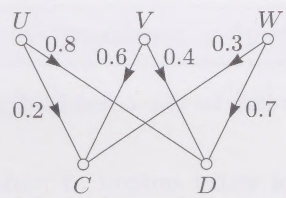


Figure 1.2 Network diagram with three input nodes and two output nodes

(a) How much water is output at nodes C and D if 1 litre of water is input at node U and 2 litres of water are input at node V? (There is no input at node W.)

(b) How much water is output at nodes C and D if x litres of water are input at node U, y litres of water are input at node V, and z litres of water are input at node W?

Solutions are given on page 46.

At the end of the video sequence you saw the effect of putting together the two networks represented in Figures 1.1 and 1.2. This combined network is shown in Figure 1.3.

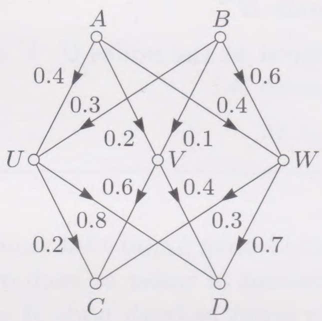


Figure 1.3 The combined network shown in the video

Using this diagram of the combined network, the amount of water flowing from the input nodes to the output nodes can be calculated. For example, suppose that 1 litre of water is input at node A (and there are no other inputs). Then we know from the discussion preceding Activity 1.1 that the amounts of water passing through nodes U, V and W are 0.4, 0.2 and 0.4 litres, respectively. These 'outputs' from the top network can be viewed as 'inputs' to the bottom network. To show how this is done, consider the output node C. From the network diagram it can be seen that:

◇ 0.2 of the water passing through node U reaches node C, so $0.2 \times 0.4 = 0.08$ litres of water from node A reach node C via node U;

◇ 0.6 of the water passing through node V reaches node C, so $0.6 \times 0.2 = 0.12$ litres of water from node A reach node C via node V;

◇ 0.3 of the water passing through node W reaches node C, so $0.3 \times 0.4 = 0.12$ litres of water from node A reach node C via node W.

To find the total amount of water output at node C, we add the contributions from the three different routes from A, to obtain $0.08 + 0.12 + 0.12 = 0.32$ litres of water output at node C.

Activity 1.4 Flow through a combined network

Both parts of this activity refer to the combined network as shown in Figure 1.3.

(a) Calculate the amount of water output at node D if the only input is 1 litre of water at node A.

(b) Calculate the amount of water output at nodes C and D if the only input is 1 litre of water at node B.

Solutions are given on page 46.

The combined network can be considered as equivalent to a simpler network with two inputs at nodes A and B and two outputs at nodes C and D. The above calculations would be a lot simpler if we could find the network diagram for this simpler network. In fact, that is what we have just done. Remember that in the video, when numbers were first associated with pipes, the number that labelled a pipe was the amount of water that flowed down the pipe when 1 litre of water was input to the node at the top.

From the calculations preceding Activity 1.4, if 1 litre of water is input at node A, then 0.32 litres of water will be output at node C. This means that the pipe from node A to node C in the simpler network can be labelled by the number 0.32. The calculations in Activity 1.4 provide the labels for the other three pipes, giving the network diagram shown in Figure 1.4.

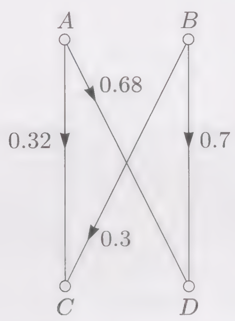


Figure 1.4 A two-input, two-output network which is equivalent to the combined network shown in Figure 1.3

To see how much easier the calculations are with this new network, try the following activity.

Activity 1.5 Flow through a two-input, two-output network

All parts of this activity refer to the two-input, two-output network shown in Figure 1.4.

(a) Calculate the amount of water output at nodes C and D if the only input is 1 litre of water at node B.

(b) Calculate the amount of water output at nodes C and D if x litres of water are input at node A and y litres of water are input at node B.

Solutions are given on page 47.

1.2 From networks to matrices

In Subsection 1.1 you calculated the amount of water leaving a pipe network when various quantities of water were input. The calculations were lengthy to write out because you constantly needed to write phrases such as 'input at node A' or 'output at node U' or 'amount reaching node C via node U'. Mathematicians generally try to avoid superfluous words, so these calculations are usually expressed more succinctly using *matrices*, which are introduced in this subsection.

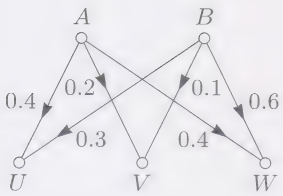


Figure 1.5 The first network seen in the video

We start by considering again our first network, shown in Figure 1.5. This network has two inputs and three outputs. The amount of water input at the two input nodes can be represented by a column of two numbers. For example, the column $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ represents an input of 1 litre of water at node A and 0 litres of water at node B. So we have replaced the phrase 'litres of water input at node A' by a specified position in a column of numbers by ordering the nodes (in this case alphabetical order). The column of numbers above is called a **vector**, and the numbers appearing in the column are called its **components**. In the above example, 1 is called the first component and 0 is called the second component.

Here are some more examples of vectors.

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \begin{pmatrix} x \\ y \end{pmatrix} \quad \begin{pmatrix} 0.4 \\ 0.2 \\ 0.4 \end{pmatrix}$$

Vectors will be used to represent other quantities of interest in later sections.

In the context of the network in Figure 1.5, the first of these vectors represents 0 litres of water input at node A and 1 litre of water input at node B. The second vector represents 1 litre of water input at node A and 2 litres of water input at node B. The third vector is the general case of x litres of water input at node A and y litres of water input at node B. The fourth vector could be used to represent the outputs from the network – namely, 0.4 litres of water output at node U, 0.2 litres of water output at node V, and 0.4 litres of water output at node W – corresponding to some vector of inputs.

Try the following activity, which reinforces the link between vectors and the inputs to and outputs from networks.

Activity 1.6 Representing inputs and outputs by vectors

All parts of this activity refer to the first network that we considered, as shown in Figure 1.5.

(a) How can inputs of 2 litres of water at node A and 3 litres of water at node B be represented by a vector?

(b) What is the first component of the vector that you wrote down as your answer to part (a)?

(c) What outputs from the network are represented by the vector $\begin{pmatrix} 0.3 \\ 0.1 \\ 0.6 \end{pmatrix}$?

Solutions are given on page 47.

Consequently, we shall refer to 'input' and 'output' vectors.

Having seen that the inputs to and outputs from a network can be represented by vectors, we turn our attention to representing the network itself. How do we represent the way in which the outputs are related to the inputs? The answer is to use a matrix, so we now proceed to define this concept.

A **matrix** is a rectangular array of numbers, which can be used to hold quantitative information in a structured way. Examples are

$$\begin{pmatrix} 1 & 2 & -1 \\ 5 & 3 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

The word 'matrix' is usually pronounced 'may-trix'.

Each number in a matrix is called an **element** of that matrix. So, for example, the first matrix above has six elements and the second matrix has four elements.

Note that although one sometimes talks about the elements of a vector, one never talks of the components of a matrix.

The elements of a matrix are arranged in rows and columns. A **row** of a matrix is a horizontal line of numbers in the matrix. For example, $(1 \quad 2 \quad -1)$ is the first row of the first matrix written above, and $(5 \quad 3 \quad 0)$ is the second row of that matrix. A **column** of a matrix is a vertical line of numbers in the matrix. For example, the first matrix written above has three columns, which are $\begin{pmatrix} 1 \\ 5 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

There is a similarity between the columns of a matrix as shown above and the vectors defined earlier: this is no coincidence. In this chapter, 'vector' is simply a shorter name for 'matrix with one column'. In the next chapter, you will meet vectors in a different context.

In order to represent a network by a matrix, the numbers associated with each pipe should appear in the matrix. There are several ways of doing this. By convention, the network in Figure 1.5 is represented by the matrix

$$\begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix}.$$

You can see that the elements in a column specify the flow *from* a node (A or B), and the elements in a row specify the flow *to* a node (U, V or W). The reason for adopting this convention will be apparent in the discussion following Activity 1.7.

Now see if you can apply this convention to represent the second network that you saw in the video.

Activity 1.7 Representing a network

The second network that you saw in the video is shown in Figure 1.6.

(a) Write down a vector which represents an input of 1 litre of water at node V, and no other inputs.

(b) How many components does the input vector have? How many columns do you expect the matrix representing the network to have?

(c) How many components does the output vector have? How many rows do you expect the matrix representing the network to have?

(d) Using your answers to parts (b) and (c), write down the matrix which represents this network.

Solutions are given on page 47.

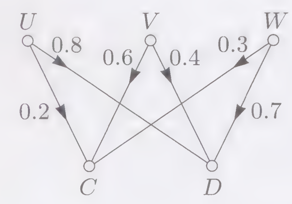


Figure 1.6 The second network seen in the video

The following discussion relates to the network in Figure 1.5, which was represented by the matrix

$$\begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix}.$$

Now we look at some of the results calculated earlier for networks, and see how they translate into matrix notation. The first thing we did was to calculate the outputs from the network in Figure 1.5 when 1 litre of water was input at node A. The inputs are represented by the vector

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The outputs calculated were 0.4 litres of water at node U, 0.2 litres of water at node V, and 0.4 litres of water at node W, which can be expressed in vector form as

$$\begin{pmatrix} 0.4 \\ 0.2 \\ 0.4 \end{pmatrix}.$$

How do we represent the relationship between the input vector and the output vector, using matrices? We take our cue from the way that functions are represented. For example, if a *function* f takes an input *number*, x, to an output *number*, y, then this can be written as $f(x) = y$. By analogy, we represent the relationship between the input *vector* and the output *vector* as

$$\begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.2 \\ 0.4 \end{pmatrix}.$$

The interpretation of this equation is that the input vector in the middle is *acted upon* by the matrix to give the output vector on the right.

The above equation was derived from results calculated using network diagrams. What is the rule for calculating the output vector given the matrix of a network and an input vector? This is what we now proceed to find.

In Activity 1.2 you considered the general case of x litres of water input at node A and y litres of water input at node B, and found that

$0.4x + 0.3y$ litres of water are output at node U,

$0.2x + 0.1y$ litres of water are output at node V,

$0.4x + 0.6y$ litres of water are output at node W.

This corresponds to the output vector

$$\begin{pmatrix} 0.4x + 0.3y \\ 0.2x + 0.1y \\ 0.4x + 0.6y \end{pmatrix}.$$

So, in matrix notation, the relationship between the input vector and the output vector is

$$\begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.4x + 0.3y \\ 0.2x + 0.1y \\ 0.4x + 0.6y \end{pmatrix}.$$

Now look at this equation. The *first* component of the output vector is the sum of the products of each of the numbers in the *first* row of the matrix with the corresponding component in the input vector, as highlighted in Figure 1.7.

$$\begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.4x + 0.3y \\ 0.2x + 0.1y \\ 0.4x + 0.6y \end{pmatrix}$$

Figure 1.7 Obtaining the first component of the output vector

Similarly, the second and third components of the output vector are obtained from the second and third rows of the matrix, respectively, as shown in Figure 1.8.

$$\begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.4x + 0.3y \\ 0.2x + 0.1y \\ 0.4x + 0.6y \end{pmatrix} \qquad \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.4x + 0.3y \\ 0.2x + 0.1y \\ 0.4x + 0.6y \end{pmatrix}$$

Figure 1.8 Obtaining the second and third components of the output vector

So we can find any component of the output vector from the elements in the corresponding row of the matrix and the components of the input vector, by forming the sum of their products. This operation is called **matrix–vector multiplication**. Note that a multiplication sign is not used to denote this operation; the matrix and vector are simply written next to each other.

This subsection ends with activities which ask you to find the output from a network when given the matrix representing the network, instead of the network diagram. The first activity refers to a familiar network for which you have the network diagram to guide you; the second activity refers to a new network.

Activity 1.8 Using matrices to calculate outputs

The second network that you saw in the video is shown in Figure 1.6.

(a) Write down a vector which represents an input of 1 litre of water at node U, and no other inputs.

(b) The matrix representing the network was found in Activity 1.7. Use this matrix to calculate the output vector for the above input vector, and interpret your answer in terms of the amounts of water collected at nodes C and D.

Solutions are given on page 47.

Activity 1.9 Finding outputs

A network has two input nodes and two output nodes. Its matrix is

$$\begin{pmatrix} 0.33 & 0.25 \\ 0.67 & 0.75 \end{pmatrix}.$$

Find the output vector for each of the following input vectors.

(a) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (b) $\begin{pmatrix} 10 \\ 14 \end{pmatrix}$ (c) $\begin{pmatrix} 18 \\ 6 \end{pmatrix}$

Solutions are given on page 47.

Matrices were invented by Arthur Cayley (1821–95). In 1858, in a paper presented to the Royal Society, he gave the rules for matrix operations (covered in Section 2) and the conditions under which a matrix has an inverse (covered in Section 5). However, it was Cayley's friend and fellow mathematician James Joseph Sylvester (1814–97) who coined the term 'matrix', in 1850.

1.3 From combining networks to matrix multiplication

In Subsection 1.1 you saw that two networks can be combined to form a simpler equivalent network. To be more precise, you saw how the first two networks on the video can be combined to produce the two-input, two-output network shown in Figure 1.9.

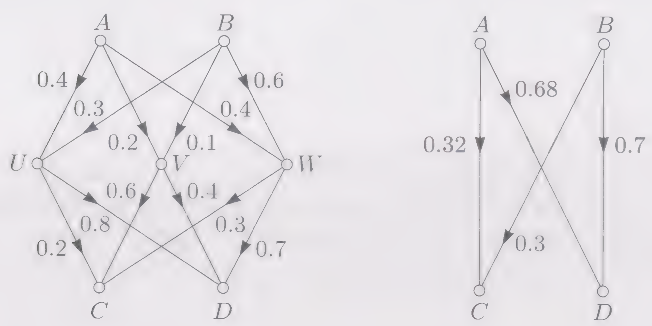


Figure 1.9 Two equivalent networks

It turns out that this simplification can also be achieved by 'combining' matrices. Suppose that the input to the left-hand network shown in Figure 1.9 is the vector $\begin{pmatrix} x \\ y \end{pmatrix}$. Then the output from the top part of the network can be written in the form

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.1}$$

This three-component output vector is used as the input to the bottom part of the left-hand network. Thus the output from the network is given by the following matrix–vector multiplication:

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}.$$

We can substitute for the three-component vector using equation (1.1), to obtain the output from the whole network in terms of the input:

This expression for the output must evaluate as a vector with two components, but we do not yet know how the two matrices on the left combine.

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.2}$$

Now consider the right-hand network shown in Figure 1.9, which is represented by the matrix

$$\begin{pmatrix} 0.32 & 0.3 \\ 0.68 & 0.7 \end{pmatrix}.$$

The output from the network for the input $\begin{pmatrix} x \\ y \end{pmatrix}$ is

$$\begin{pmatrix} 0.32 & 0.3 \\ 0.68 & 0.7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.3}$$

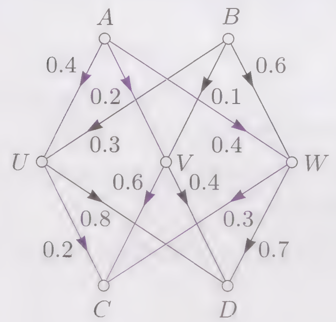
Since the two networks are equivalent, expression (1.2) must be equal to expression (1.3). This will happen if there is a sense in which meaning can be given to the left-hand side of the matrix equation

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} = \begin{pmatrix} 0.32 & 0.3 \\ 0.68 & 0.7 \end{pmatrix}. \tag{1.4}$$

The notion of equality between matrices is defined in Section 2.

Equation (1.4) was deduced by considering the combination and simplification of two networks. We now define a method of 'combining' matrices that is consistent with equation (1.4). This method of 'combining' matrices is called *matrix multiplication*, and the 'combined' matrix is called the *product matrix*.

We start by considering the element 0.32 in the top left-hand corner of the product matrix $\begin{pmatrix} 0.32 & 0.3 \\ 0.68 & 0.7 \end{pmatrix}$. How was this calculated? If you look back to Subsection 1.1 (page 8), then you will see that we added the contributions flowing down each pair of pipes that could lead from node A to node C. This is highlighted in the left-hand diagram of Figure 1.10. In the right-hand diagram of Figure 1.10, the corresponding matrix elements are highlighted.



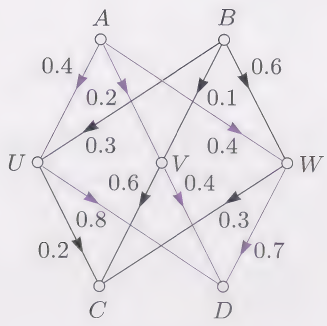
$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix}$$

Figure 1.10 Highlighting shows pipes and corresponding matrix elements

Looking at the right-hand diagram in Figure 1.10, we see that the *top left* entry of the product matrix is formed from the *top* row of the first matrix and the *left*-hand column of the second matrix. What is the rule for combining a row and a column? The answer is exactly the same as when calculating a matrix–vector product: add together the products of corresponding elements. In this case, we obtain

$$(0.2 \times 0.4) + (0.6 \times 0.2) + (0.3 \times 0.4) = 0.32.$$

Now let us see this idea at work again. How is the bottom left element, 0.68, of the product matrix formed? In terms of combining networks, it is the sum of the contributions flowing from node A to node D via any of the three possible routes. In terms of the matrix product, we expect that the *bottom left* element of the product matrix should be formed from the *bottom* row of the first matrix and the *left*-hand column of the second matrix. The network routes and matrix elements are highlighted in Figure 1.11, overleaf.



$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix}$$

Figure 1.11 Highlighting shows pipes and corresponding matrix elements

From the elements highlighted in the matrix multiplication above, we can calculate the corresponding element of the matrix product:

$$(0.8 \times 0.4) + (0.4 \times 0.2) + (0.7 \times 0.4) = 0.68.$$

Matrix manipulation is the subject of Section 2.

The process of combining two matrices as described above is known as **matrix multiplication**. The result of multiplying two matrices is called the **product matrix**.

The next activity asks you to verify that the other two elements in the product matrix $\begin{pmatrix} 0.32 & 0.3 \\ 0.68 & 0.7 \end{pmatrix}$ are correct.

Activity 1.10 Combining networks and multiplying matrices

(a) Consider the top right element, 0.3, in the above product matrix.

(i) Sketch the network diagram, and highlight the routes through the network that contribute to this element of the product matrix.

(ii) Write out the matrix multiplication, and highlight the elements of the matrices that contribute to this element.

(iii) Use the highlighted elements of the matrices to re-calculate the element of the product matrix.

(b) Repeat part (a) for the the bottom right element of the product matrix.

Solutions are given on page 48.

This section ends with another look at finding outputs using matrices and multiplying matrices on the video.

Now watch the video to the end of Band B.

Summary of Section 1

This section has introduced:

- ◇ the idea of a network, and how to calculate the output from a network and from combined networks;
- ◇ vectors, as a notation for inputs to and outputs from a network;
- ◇ matrices, as a notation for representing networks;
- ◇ matrix–vector multiplication, as a way of calculating the output from a network;
- ◇ matrix multiplication, by considering the combination of networks.

Exercise for Section 1

Exercise 1.1

(a) Find the output vector when each of the following input vectors is acted on by the matrix $\begin{pmatrix} 0.35 & 0.85 \\ 0.65 & 0.15 \end{pmatrix}$.

(i) $\begin{pmatrix} 2 \\ 6 \end{pmatrix}$ (ii) $\begin{pmatrix} 50 \\ 100 \end{pmatrix}$ (iii) $\begin{pmatrix} 60 \\ 40 \end{pmatrix}$

(b) Sketch the network corresponding to the matrix in part (a), including the appropriate label for each pipe.

2 Matrix manipulation

This section aims to help strengthen your understanding of matrix multiplication and introduces some other ways of manipulating matrices.

2.1 Matrix multiplication

Section 1 introduced matrix multiplication and in particular evaluated the matrix product

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix} = \begin{pmatrix} 0.32 & 0.3 \\ 0.68 & 0.7 \end{pmatrix}. \tag{2.1}$$

There is something special about the numbers of rows and columns of the matrices appearing in equation (2.1). In order to discuss this, we need some notation. The left-hand matrix has 2 rows and 3 columns. We often abbreviate this and say that the left-hand matrix is a 2×3 matrix. The next matrix has 3 rows and 2 columns, so it is a 3×2 matrix.

'2×3' is read as 'two by three'.

In general, an $\boldsymbol{m \times n}$ **matrix** is one with m rows and n columns. We say that the **size** of the matrix is $m \times n$. Also, two matrices of the same size are **equal** if all their corresponding elements agree.

Some people prefer to say *shape* or *order* rather than *size* here.

The special nature of the sizes of the matrices in the above matrix product is best seen by writing the sizes next to each other as follows:

$$2 \times 3 \quad 3 \times 2 \quad \text{gives} \quad 2 \times 2.$$

The first observation is that the two 3s are adjacent to each other. There is a good reason for this: to combine a row and a column, they must each have the same number of elements. The first matrix has 3 elements in each row and the second matrix has 3 elements in each column.

We can indicate the matching of the 3s by drawing a box around them:

$$2 \times \boxed{3 \quad 3} \times 2 \quad \text{gives} \quad 2 \times 2.$$

If we now ignore the numbers in the box, then on the left we have the numbers 2 and 2 which appear in the size of the product matrix, 2×2.

In fact, this always happens. Two matrices can be multiplied only if the adjacent numbers match. Furthermore, the size of the product matrix is given by the remaining numbers when we ignore the adjacent numbers.

Thus we have a useful method for deciding whether two matrices can be multiplied and for obtaining the size of a product matrix. Try the following activity to practise this method.

Activity 2.1 Determining the size of matrix products

For each of the following, state whether the two matrices can be multiplied. If they can, then determine the size of the product matrix. (Do not attempt to calculate the product matrix.)

(a) $\begin{pmatrix} \frac{1}{2} & 9 \\ 7 & \frac{1}{3} \\ -6 & 8 \end{pmatrix} \begin{pmatrix} -2 & 3 & 1 \\ 4 & 0 & 5 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 9 \\ 8 \\ 7 \end{pmatrix}$

(c) $\begin{pmatrix} \frac{1}{2} & 9 \\ 7 & \frac{1}{3} \\ -6 & 8 \end{pmatrix} \begin{pmatrix} 0 & 10 \\ 4 & \frac{1}{2} \\ -2 & 7 \end{pmatrix}$ (d) $\begin{pmatrix} 9 \\ 8 \\ 7 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 2 & 3 \\ 7 & 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 2 & 1 & 3 \end{pmatrix}$

Solutions are given on page 48.

Before continuing, here is the general definition of matrix multiplication, which summarises what we have found above.

Matrix multiplication

Let **A** and **B** be two matrices. We can form the product matrix **AB** only if the number of columns of **A** equals the number of rows of **B**.

The product matrix **AB** is formed by 'combining' rows of matrix **A** with columns of matrix **B**, where 'combining' means adding the products of corresponding elements. The rule for where to put the resulting numbers is that the element in the ith row and jth column of **AB** is the result of 'combining' the ith row of **A** with the jth column of **B**.

Note that we use the convention of upper-case bold letters to denote matrices. When writing these letters by hand, they are usually written as plain upper-case letters (with no underlining or other adornments).

Several ways have been developed to help apply the above definition. If you are good at mental arithmetic, one of the simplest ways is to lay your hands along the rows and columns of the matrices as you are multiplying them. For example, to calculate the top left element of the product matrix, place your hands as shown in Figure 2.1. (If you need to use pen and paper or a calculator, slips of paper can play the same role as the hands.)

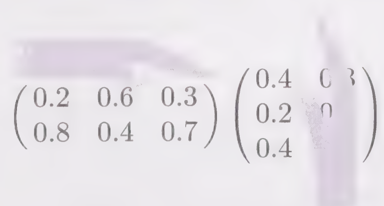


Figure 2.1 Laying hands on a matrix multiplication to outline rows and columns

The corresponding elements can then be read off *underneath* the left hand and on the *left* of the right hand; in this case we obtain

$$0.2 \times 0.4 + 0.6 \times 0.2 + 0.3 \times 0.4.$$

The next activity provides some practice in matrix multiplication.

Activity 2.2 Calculating matrix products

For each of the following pairs of matrices $\mathbf{A}$ and $\mathbf{B}$, calculate the matrix product $\mathbf{AB}$.

For which of these pairs of matrices can you also find the matrix product $\mathbf{BA}$? Calculate those which can be formed. For those which cannot be formed, say why this is so.

(a) $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \\ 6 & 1 \end{pmatrix}$, $\quad \mathbf{B} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & -1 & 1 \end{pmatrix}$

(b) $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$, $\quad \mathbf{B} = \begin{pmatrix} 4 & 6 \\ 2 & -1 \end{pmatrix}$

(c) $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \\ 4 & 6 \end{pmatrix}$, $\quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

Solutions are given on page 49.

This way of combining matrices is called matrix multiplication because it shares many of the properties of multiplication of numbers. However, there is one important *difference* between the properties for matrices and for real numbers. The order in which you multiply two real numbers does not matter, since $ab = ba$ for all real numbers a and b. This is *not* true for matrix multiplication. There are matrices $\mathbf{A}$ and $\mathbf{B}$ for which the product matrices $\mathbf{AB}$ and $\mathbf{BA}$ can be formed, but for which

$$\mathbf{AB} \neq \mathbf{BA}.$$

If you have not noticed this already, then look back at Activity 2.2(b), which shows that even if the product matrices are the same size, $\mathbf{AB}$ might not be equal to $\mathbf{BA}$.

Powers of matrices

Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}.$$

Just as x^2 is defined as the product $x \times x$, so we can define the square of $\mathbf{A}$, written $\mathbf{A}^2$, as the matrix product $\mathbf{AA}$; thus

Note that $\mathbf{A}^2$ is the same size as $\mathbf{A}$.

$$\mathbf{A}^2 = \mathbf{AA} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 7 & 2 \\ 6 & 3 \end{pmatrix}.$$

Not all matrices can be squared, however. The next activity considers the possibility of squaring a matrix.

Activity 2.3 When does the square of a matrix exist?

(a) Let $\mathbf{M} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 4 \end{pmatrix}$.

Why is it not possible to form the matrix product $\mathbf{MM}$?

(b) Use your answer to part (a) to formulate the condition on a matrix $\mathbf{A}$ under which it is possible to calculate its square $\mathbf{A}^2$.

Solutions are given on page 49.

Any matrix with the same number of rows as columns is called a **square matrix**. So $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$ is square, but $\mathbf{M} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 4 \end{pmatrix}$ is not. From the above activity, the condition for the square of a matrix $\mathbf{A}$ to exist is that $\mathbf{A}$ is a square matrix.

We can also form the products $(\mathbf{A}^2)\mathbf{A}$ and $\mathbf{A}(\mathbf{A}^2)$, because $\mathbf{A}$ is a square matrix and $\mathbf{A}^2$ is a matrix of the same size as $\mathbf{A}$. For example,

Above, we found

$$\mathbf{A}^2 = \begin{pmatrix} 7 & 2 \\ 6 & 3 \end{pmatrix}.$$

$$(\mathbf{A}^2)\mathbf{A} = \begin{pmatrix} 7 & 2 \\ 6 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 20 & 7 \\ 21 & 6 \end{pmatrix},$$

$$\mathbf{A}(\mathbf{A}^2) = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 7 & 2 \\ 6 & 3 \end{pmatrix} = \begin{pmatrix} 20 & 7 \\ 21 & 6 \end{pmatrix}.$$

In this example $(\mathbf{A}^2)\mathbf{A} = \mathbf{A}(\mathbf{A}^2)$, so we can write this product as $\mathbf{AAA}$, without ambiguity. We denote $\mathbf{AAA}$ by $\mathbf{A}^3$; that is, the cube of $\mathbf{A}$ is

$$\mathbf{A}^3 = \mathbf{AAA}.$$

For any positive integer n, the ***n*th power** of a square matrix $\mathbf{A}$, written $\mathbf{A}^n$, is defined in the same way, by repeated multiplication.

The cube of a matrix is always unambiguously defined because it does not matter which matrix multiplication is done first. In fact, more generally, it is true that

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

This equation means that the matrix product $\mathbf{ABC}$ is unambiguous.

for any matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, in the sense that if the matrix product on either side of the equation exists, then so does the other and they are equal. One way of visualising this fact is to think about combining three networks. For example, consider the three networks shown in Figure 2.2, where network 1 is represented by the matrix $\mathbf{C}$, network 2 is represented by the matrix $\mathbf{B}$, and network 3 is represented by the matrix $\mathbf{A}$. You can think of the entire network as network 1 and network 2 combined ($\mathbf{BC}$), and then combining this with network 3 ($\mathbf{A}(\mathbf{BC})$). Or you can think of it as network 2 and network 3 combined ($\mathbf{AB}$), and then network 1 added on top ($(\mathbf{AB})\mathbf{C}$). But changing the way you 'think' about the entire network does not change the physical network, so the flow through it is the same.

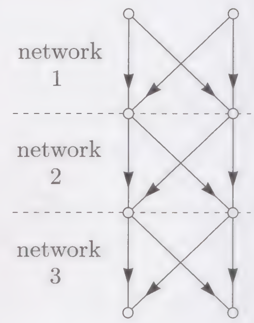


Figure 2.2 Combining three networks

2.2 Matrix addition

Until now, all our efforts have been devoted to the motivation for, and mechanics of, multiplying matrices. This is because the importance of matrices arises from matrix multiplication. But matrices can also be added, and this is sometimes useful. (For example, vector addition has an important geometric interpretation.)

The geometric approach to vector addition is discussed in Chapter B3.

Now **matrix addition** is introduced. Two matrices can be added only if they have the same size. To add two matrices, just add the corresponding elements. For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1+7 & 2+8 & 3+9 \\ 4+10 & 5+11 & 6+12 \end{pmatrix}$$
$$= \begin{pmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix}.$$

The rule for adding matrices extends to the sum of any number of matrices of the same size: just add corresponding elements. In particular, this means that

$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}),$$

so the expression $\mathbf{A}+\mathbf{B}+\mathbf{C}$ is unambiguous.

Now try matrix addition yourself.

Activity 2.4 Adding matrices

For each of the pairs of matrices below, calculate the matrix sums $\mathbf{A}+\mathbf{B}$ and $\mathbf{B}+\mathbf{A}$.

(a) $\mathbf{A}=\begin{pmatrix}2 & 1\\ 3 & 0\end{pmatrix}$, $\mathbf{B}=\begin{pmatrix}4 & 6\\ 2 & -1\end{pmatrix}$

(b) $\mathbf{A}=\begin{pmatrix}2 & 0 & 4\\ -1 & 6 & 0\end{pmatrix}$, $\mathbf{B}=\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}$

(c) $\mathbf{A}=\begin{pmatrix}4 & 0 & 5\\ 6 & 1 & 0\\ 2 & 3 & -1\end{pmatrix}$, $\mathbf{B}=\begin{pmatrix}6 & 2 & 0\\ 0 & -1 & 4\\ 5 & 0 & 1\end{pmatrix}$

Solutions are given on page 49.

Activity 2.4 demonstrates that order is unimportant in matrix addition; that is, $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$ for all matrices $\mathbf{A}$ and $\mathbf{B}$ of the same size.

Matrix addition has an important special case, which is **vector addition**: the sum of two vectors with the same number of components is formed by adding corresponding components. For example,

$$\begin{pmatrix}3\\ 0\\ -1\end{pmatrix}+\begin{pmatrix}5\\ -3\\ 7\end{pmatrix}=\begin{pmatrix}3+5\\ 0+(-3)\\ -1+7\end{pmatrix}=\begin{pmatrix}8\\ -3\\ 6\end{pmatrix}.$$

Now try adding vectors yourself.

Activity 2.5 Adding vectors

(a) Find the following vector sums.

(i) $\begin{pmatrix}0.83\\ 0.17\end{pmatrix}+\begin{pmatrix}0.36\\ 0.64\end{pmatrix}$

(ii) $\begin{pmatrix}-5\\ 7\\ -4\end{pmatrix}+\begin{pmatrix}-3\\ 1\\ 4\end{pmatrix}$

(iii) $\begin{pmatrix}6\\ 1\end{pmatrix}+\begin{pmatrix}-2\\ 2\end{pmatrix}+\begin{pmatrix}9\\ 3\end{pmatrix}$

(b) Give an example of a pair of vectors which cannot be added together.

Solutions are given on page 49.

Scalar multiplication of matrices

Any matrix can be added to itself. Often $\mathbf{A}+\mathbf{A}$ is written as $2\mathbf{A}$. This notation can be generalised to any real number k and any matrix $\mathbf{A}$, so that $k\mathbf{A}$ is the matrix with each element equal to k times the corresponding element in $\mathbf{A}$. For example,

$$\tfrac{1}{2}\begin{pmatrix} 4 & 3 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\times 4 & \frac{1}{2}\times 3 \\ \frac{1}{2}\times 0 & \frac{1}{2}\times(-1) \end{pmatrix} = \begin{pmatrix} 2 & \frac{3}{2} \\ 0 & -\frac{1}{2} \end{pmatrix}.$$

The matrix $k\mathbf{A}$ is called the **scalar multiple** of the matrix $\mathbf{A}$ by the real number k, and the operation of forming $k\mathbf{A}$ is called **scalar multiplication** of matrices.

In this context, the real number k is often called a **scalar**.

Now try applying this operation for yourself.

Activity 2.6 Scalar multiplication of matrices

(a) In each case below, simplify the expression by performing the scalar multiplication.

(i) $3\begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}$ (ii) $\frac{1}{2}\begin{pmatrix} 10 & -3 \\ 7 & 3 \end{pmatrix}$ (iii) $\frac{2}{3}\begin{pmatrix} 6 & x \\ 3 & y \end{pmatrix}$

(b) Simplify each of the following matrices by writing it as a scalar multiple of a matrix with integer entries.

(i) $\begin{pmatrix} 4.5 & 3 \\ 2 & 1.5 \end{pmatrix}$ (ii) $\begin{pmatrix} \frac{3}{2} \\ \frac{2}{3} \end{pmatrix}$

(iii) $\begin{pmatrix} 2x & 3x \\ 0 & 5x \end{pmatrix}$, where x is a real number.

(c) Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$. Show that

$$\mathbf{A}(2\mathbf{B}) = 2(\mathbf{AB}).$$

Solutions are given on page 50.

Activity 2.6(c) illustrates the following useful result concerning scalar multiplication. If the matrix product $\mathbf{AB}$ exists, then

$$\mathbf{A}(k\mathbf{B}) = (k\mathbf{A})\mathbf{B} = k(\mathbf{AB}), \tag{2.2}$$

where k is a real number.

The operation of **matrix subtraction** is defined in terms of scalar multiplication and addition of matrices. The result of subtracting a matrix $\mathbf{B}$ from a matrix $\mathbf{A}$, written $\mathbf{A}-\mathbf{B}$, is given by

$$\mathbf{A}-\mathbf{B} = \mathbf{A}+(-1)\mathbf{B},$$

where $(-1)\mathbf{B}$ is the scalar multiple of $\mathbf{B}$ by -1. Thus, only matrices of the same size can be subtracted, and the subtraction is performed by subtracting corresponding elements. The following example illustrates this operation:

The matrix $(-1)\mathbf{B}$ is often written as $-\mathbf{B}$, and is referred to as the 'negative of $\mathbf{B}$'.

$$\begin{pmatrix} 4 & 5 \\ -2 & 7 \end{pmatrix} - \begin{pmatrix} 3 & 9 \\ 8 & -1 \end{pmatrix} = \begin{pmatrix} 4-3 & 5-9 \\ -2-8 & 7-(-1) \end{pmatrix} = \begin{pmatrix} 1 & -4 \\ -10 & 8 \end{pmatrix}.$$

Factorising

When solving problems involving matrices, you may encounter a matrix expression of the form

$$\mathbf{AB} + \mathbf{AC}.$$

It can be shown that this expression can be factorised, to give

$$\mathbf{AB} + \mathbf{AC} = \mathbf{A}(\mathbf{B} + \mathbf{C}). \tag{2.3}$$

This equation holds whenever the matrix products and sums on each side exist.

Activity 2.7 asks you to verify equation (2.3) in a particular case, where **B** and **C** are vectors.

Activity 2.7 Verifying a result

Let $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 3 & 4 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{C} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Show that

$$\mathbf{AB} + \mathbf{AC} = \mathbf{A}(\mathbf{B} + \mathbf{C}).$$

A solution is given on page 50.

Element notation

Sometimes it is useful to have a notation for the elements of a matrix. Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

This is a 2×2 matrix, and its elements can be described in terms of row and column positions using subscripts:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

A comma can be added if a subscript is ambiguous. For example, a_{112} could be written as $a_{11,2}$ or $a_{1,12}$, as appropriate.

Note that the lower-case italic letter a has been used to denote the elements of the matrix **A**, to emphasise that they are numbers rather than matrices. The subscript is used to specify the position of the element, as follows:

a_{11} is the element in the first row and the first column ($a_{11} = 1$);
a_{12} is the element in the first row and the second column ($a_{12} = 2$);
a_{21} is the element in the second row and the first column ($a_{21} = 3$);
a_{22} is the element in the second row and the second column ($a_{22} = 4$).

The following diagram might help you to remember this.

$$\begin{array}{cc} & \begin{array}{cc} \text{col 1} & \text{col 2} \\ \downarrow & \downarrow \end{array} \\ \begin{array}{c} \text{row 1} \rightarrow \\ \text{row 2} \rightarrow \end{array} & \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \end{array}$$

The following activity tests your understanding of this notation.

Activity 2.8 Interpreting element notation

For the matrix

$$\mathbf{B} = \begin{pmatrix} 7 & 2.5 & 0 \\ -4 & 1 & 9 \end{pmatrix},$$

write down the following elements.

(a) b_{11} (b) b_{21} (c) b_{13}

Solutions are given on page 50.

As an example of how element notation can be used, here is a succinct statement of the rule for matrix addition.

Matrix addition

If $\mathbf{C}$ is the sum of two $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, then $\mathbf{C}$ is an $m \times n$ matrix with elements

$$c_{ij} = a_{ij} + b_{ij},$$

where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Element notation can also be used to state concisely the rule for matrix multiplication. To see how to do this, consider two 2×2 matrices $\mathbf{A}$ and $\mathbf{B}$, and let $\mathbf{AB} = \mathbf{C}$. We can write

$$\mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \mathbf{C}.$$

Now consider the top left element c_{11} of the product matrix. This element is formed by adding the products of the corresponding elements in the top row of $\mathbf{A}$ and the left-hand column of $\mathbf{B}$:

$$c_{11} = (a_{11} \times b_{11}) + (a_{12} \times b_{21}).$$

The same pattern of subscripts appears in the calculation of the other elements, and in general an element of the product matrix is given by

$$c_{ij} = (a_{i1} \times b_{1j}) + (a_{i2} \times b_{2j}), \text{ where } i = 1, 2, \; j = 1, 2.$$

You can check this by multiplying out the above matrix product. This formula applies to any 2×2 matrix (or indeed any matrix product where the first matrix has two columns and the second matrix has two rows, for example, when multiplying a 3×2 matrix by a 2×3 matrix).

For multiplying two 3×3 matrices, the rule is

$$c_{ij} = (a_{i1} \times b_{1j}) + (a_{i2} \times b_{2j}) + (a_{i3} \times b_{3j}), \text{ where } i = 1, 2, 3, \; j = 1, 2, 3.$$

This can be written, using the sigma notation for sums that was introduced in Chapter B1, as

See Chapter B1, Subsection 1.2.

$$c_{ij} = \sum_{k=1}^{3} a_{ik} \times b_{kj}.$$

From this 3×3 case, we generalise to the case of multiplying any two matrices of appropriate sizes.

Matrix multiplication

If $\mathbf{C}$ is the product of an $m \times n$ matrix $\mathbf{A}$ and an $n \times p$ matrix $\mathbf{B}$, then $\mathbf{C}$ is an $m \times p$ matrix with elements

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \times b_{kj},$$

where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, p$.

Note that lower-case bold symbols are used for vectors. When writing vectors by hand, you should underline the symbol either with a straight line, $\underline{v}$, or with a wavy line, $\utilde{v}$.

Element notation can also be used for the components of a vector. Since vectors always have only one column, the second subscript is always 1; for example, for a vector $\mathbf{v}$ with two components we have

$$\mathbf{v} = \begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix}.$$

The above notation is correct, but cumbersome. We usually omit the second number, and write

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Summary of Section 2

This section has introduced:

◇ multiplication and addition of two suitable matrices $\mathbf{A}$ and $\mathbf{B}$, written as $\mathbf{AB}$ and $\mathbf{A} + \mathbf{B}$, respectively;

◇ scalar multiplication of a matrix $\mathbf{A}$ by a real number k, written as $k\mathbf{A}$;

◇ some properties of matrices:

$$\begin{aligned}
&(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}),\\
&\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A},\\
&(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}),\\
&\mathbf{A}(k\mathbf{B}) = (k\mathbf{A})\mathbf{B} = k(\mathbf{AB}),\\
&\mathbf{AB} + \mathbf{AC} = \mathbf{A}(\mathbf{B} + \mathbf{C});
\end{aligned}$$

◇ the notation a_{ij} for the element in the ith row and jth column of the matrix $\mathbf{A}$.

Exercises for Section 2

In Exercises 2.1 and 2.2, use the following matrices.

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 4 & 6 & 5 \\ 2 & 1 & 3 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 1 & 4 & 2 \\ 2 & 3 & 5 \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} 1 & 2 \\ 4 & 0 \\ 0 & 3 \end{pmatrix}$$

Exercise 2.1

(a) Which of the following matrix products can be formed?

$$\mathbf{AB} \quad \mathbf{CA} \quad \mathbf{CD} \quad \mathbf{DA} \quad \mathbf{DB} \quad \mathbf{A}^2 \quad \mathbf{D}^2$$

(b) Evaluate those matrix products in part (a) which can be formed.

Exercise 2.2

(a) Which of the following combinations of matrices can be formed?

$$\mathbf{A} + \mathbf{B} \quad \mathbf{B} + \mathbf{C} \quad \mathbf{C} + \mathbf{A} \quad \mathbf{C} + \mathbf{B} \quad \mathbf{D} + \mathbf{D} \quad \mathbf{A} + \mathbf{A} + \mathbf{A} \quad \mathbf{B} - \mathbf{C}$$

(b) Evaluate those combinations in part (a) which can be formed.

Exercise 2.3

Evaluate each of the following scalar multiplications.

(a) $5\begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix}$ (b) $\frac{1}{5}\begin{pmatrix} 2 & -5 \\ 10 & 3 \end{pmatrix}$ (c) $-\frac{2}{3}\begin{pmatrix} 3 & -5 \\ 4 & -9 \end{pmatrix}$

Exercise 2.4

Evaluate each of the following.

(a) $\begin{pmatrix} -3 \\ 7 \end{pmatrix} + \begin{pmatrix} 2 \\ -4 \end{pmatrix}$ (b) $\begin{pmatrix} 0.5 \\ 0.1 \\ 0.4 \end{pmatrix} - \begin{pmatrix} 0.3 \\ 0.4 \\ 0.3 \end{pmatrix}$

(c) $\begin{pmatrix} 8 \\ -2 \end{pmatrix} + \begin{pmatrix} -8 \\ 2 \end{pmatrix}$ (d) $\begin{pmatrix} 10 \\ -3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

3 A population problem

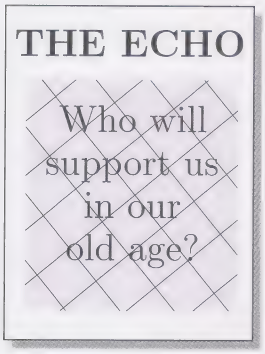


Figure 3.1 A newspaper headline

In Chapter B1 you studied the long-term behaviour of populations, using recurrence sequences to explore the ways in which a population increased or decreased – or perhaps remained the same. But many of the questions asked about populations are concerned not so much with the total population as with its structure. For example, underlying the headline question in Figure 3.1 is a concern about the proportions of pensionable and working people in the population, and how these groups may vary in the future. This suggests that a model to address this question will need to divide the total population into a number of subpopulations, and then to explore the ways in which these subpopulations behave. That is the approach adopted in this chapter.

3.1 Population modelling

In any human population, there are two major groups which are not part of the workforce: the juveniles who have not yet entered the workforce and the elderly who have retired. So to begin to investigate the question of whether the potential workforce is big enough to support the whole population, we shall divide it into three broad age groups – juveniles, workers and elderly – and investigate the proportions of each of these in the total population. Table 3.1 shows data for the human population of the UK in June 1990. (All values in the table are correct to the number of decimal places given.)

Data adapted from *Demographic Yearbook 1991*, © United Nations 1992. This book uses 15 as the age group boundary for adulthood, and 65 as the age of retirement, so we adopt these conventions too.

Table 3.1 UK population in June 1990: juveniles, workers and elderly

	Juveniles (aged < 15)	Workers (aged 15–64)	Elderly (aged ≥ 65)
Subpopulation (millions)	10.92	37.50	8.99
Birth rate	0.0000	0.0213	0.0000
Death rate	0.0008	0.0031	0.0575

The birth rates quoted in Table 3.1 are the annual proportionate birth rates, as defined in Chapter B1. For example, the number 0.0213, which we shall call the 'worker birth rate', is calculated as follows:

$$\text{worker birth rate} = \frac{\text{number of babies born with worker parents}}{\text{total number of workers}}. \qquad (3.1)$$

The death rates quoted in Table 3.1 are the annual proportionate death rates, and are calculated similarly. Test your understanding of these data by trying the following activity.

Activity 3.1 Understanding the data

(a) Use the data in Table 3.1 to estimate the number of babies born between June 1990 and June 1991.

(b) Use the data in Table 3.1 to estimate whether more workers or elderly died in the period from June 1990 to June 1991.

Solutions are given on page 50.

We have divided the population into three age groups in order to start to address the question posed in Figure 3.1. However, to establish a technique for modelling such subdivided populations, we look first at the simpler case where there are just two subpopulations – juveniles and adults. In Section 4 you will analyse the original question on the computer. The simpler question that we shall investigate in this section is the following.

> Will there be a constant proportion of juveniles entering the potential workforce, or will we have an ever-increasing proportion of adults?

The data appropriate for this simplified problem are shown in Table 3.2.

Table 3.2 UK population in June 1990: juveniles and adults

	Juveniles (aged < 15)	Adults (aged ≥ 15)
Subpopulation (millions)	10.92	46.49
Birth rate	0.0000	0.0172
Death rate	0.0008	0.0136

Note that the data in Table 3.2 are calculated directly from the data in Table 3.1. So, for example,

$$\begin{aligned}\text{adult birth rate} &= \frac{\text{total number of births to workers and to elderly}}{\text{total number of workers and elderly}}\\ &= \frac{0.0213 \times 37.50 + 0.0000 \times 8.99}{37.50 + 8.99}\\ &\simeq 0.0172.\end{aligned}$$

The death rate for adults is calculated in a similar way, from the death rates for workers and for elderly.

Now we are going to create a mathematical model to estimate the numbers of juveniles and adults in the years following 1990. The first task is to assign variables to the quantities in the model.

Table 3.2 gives figures for June 1990, and we shall take this date as the start of year 0, with year n starting n years after that date. Suppose that at the start of year n, the number of juveniles is J_n and the number of adults is A_n (in millions). In terms of these variables, A_0 is the number of millions of adults in June 1990 as given in Table 3.2, so $A_0 = 46.49$. Similarly, Table 3.2 gives $J_0 = 10.92$.

How do we calculate J_1 and A_1? In Chapter B1 the (single) population estimate for year 1 depended on only the (single) population estimate for year 0. Here, because a juvenile either becomes an adult or stays a juvenile, and because an adult stays an adult but can produce new juveniles (babies), we can expect that each of the two subpopulation estimates for year 1 will depend on *both* of the subpopulation estimates for year 0. This interdependency can be represented by the network shown in Figure 3.2. There are four pipes shown in Figure 3.2, which represent the 'flow' from one subpopulation to another. The 'quantity of flow' down each pipe can be given an interpretation in this current context, which is summarised in Table 3.3. Using this interpretation, we can determine the labels for the pipes.

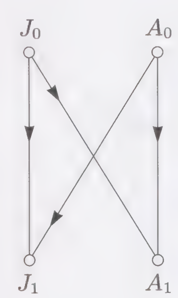


Figure 3.2 A network which shows the interrelationships between subpopulations in successive years

Table 3.3 Interpretation of the pipes in Figure 3.2

From	To	Interpretation
J_0	J_1	Proportion of juveniles who neither die nor reach age 15 in year 0
J_0	A_1	Proportion of surviving juveniles who reach age 15 in year 0
A_0	J_1	Proportionate birth rate for adults in year 0
A_0	A_1	Proportion of adults who do not die in year 0

(Note that it is convenient to use J_0, J_1, A_0 and A_1 as names for the nodes in the network as well as for estimates of subpopulation sizes.)

Consider the pipe from J_0 to J_1. There are J_0 juveniles at the start of year 0. From the data given in Table 3.2, a proportion 0.0008 of these die during the year, so a proportion $1 - 0.0008 = 0.9992$ are alive at the end of the year. Some of this number will have reached the age of 15 during the year and become adults. We shall assume that $\frac{1}{15}$ of the surviving juveniles become adults. (This is the case if there is a uniform spread of ages within the group.) So the proportion of juveniles still alive and still juveniles at the end of the year is $\frac{14}{15} \times 0.9992 = 0.9326$ to four decimal places. Thus the number that labels the pipe from J_0 to J_1 is 0.9326.

Now try to calculate the numbers that label the other pipes in the network.

Activity 3.2 Labelling the pipes

There are four pipes in the network shown in Figure 3.2, and each has an interpretation given in Table 3.3. The number labelling the pipe from J_0 to J_1 was calculated above, using the assumption that one fifteenth of surviving juveniles become adults each year. Now calculate the labels for the other pipes (correct to four decimal places).

(a) Calculate the number that labels the pipe from A_0 to J_1.

(b) Calculate the number that labels the pipe from A_0 to A_1.

(c) Calculate the number that labels the pipe from J_0 to A_1.

(*Hint*: Use the '$\frac{1}{15}$' assumption.)

Solutions are given on page 50.

The network model of the changes in the two subpopulations from year 0 to year 1 can be converted into a matrix model using the correspondence established in Subsection 1.2. So the estimates for the two subpopulations in year 1 can be calculated by

$$\begin{pmatrix} J_1 \\ A_1 \end{pmatrix} = \begin{pmatrix} 0.9326 & 0.0172 \\ 0.0666 & 0.9864 \end{pmatrix} \begin{pmatrix} J_0 \\ A_0 \end{pmatrix}.$$

If we assume in our model that the birth and death rates are constant (that is, the same each year), then the same matrix can be used to calculate the subpopulation estimates for subsequent years. Thus we have the matrix model

$$\begin{pmatrix} J_{n+1} \\ A_{n+1} \end{pmatrix} = \begin{pmatrix} 0.9326 & 0.0172 \\ 0.0666 & 0.9864 \end{pmatrix} \begin{pmatrix} J_n \\ A_n \end{pmatrix} \quad (n = 0, 1, 2, \ldots). \tag{3.2}$$

Now let us pause and consider some simplifying assumptions that were made or could have been made in arriving at these equations.

Activity 3.3 Assumptions in the model

For each assumption in the list below, state whether it was needed to create the model in equation (3.2), and comment on whether you think the assumption is reasonable.

- The birth and death rates are constant.
- No juveniles give birth.
- One fifteenth of surviving juveniles become adults each year.
- There is no emigration or immigration.

Solutions are given on page 51.

Since the birth and death rates are assumed constant in this model, the network diagram in Figure 3.2 can be generalised to year n and year $n+1$, as shown in Figure 3.3. This figure includes the calculated labels.

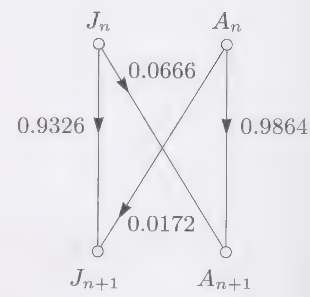


Figure 3.3
Generalising Figure 3.2

Now you are asked to use the matrix model to calculate the population estimates for year 1 and year 2.

Activity 3.4 Calculating

Use equation (3.2) and the data in Table 3.2 to estimate the values of J_1, A_1, J_2 and A_2.

State what each of these numbers represents.

Solutions are given on page 51.

By repeating the type of calculation in Activity 3.4, we obtain predictions for the seven years following 1990, shown in Table 3.4. (In this table and related ones, each prediction is for June of the given year, and $n = 0$ corresponds to 1990.)

Table 3.4 Predictions from the model (populations in millions)

	1990	1991	1992	1993	1994	1995	1996	1997
Juveniles, J_n	10.92	10.98	11.04	11.10	11.16	11.21	11.27	11.32
Adults, A_n	46.49	46.59	46.68	46.78	46.89	46.99	47.10	47.21

There is no need for you to check the predictions for 1993–1997.

Looking at the results shown in Table 3.4, it can be seen that the numbers for both subpopulations are increasing. First we investigate the increase in the *total* population, by defining T_n to be the total population at the beginning of year n, that is, $T_n = J_n + A_n$. Now the variation in the total population can be seen by calculating T_n, as shown in the first row of Table 3.5.

Table 3.5 Predictions of total population from the model

	1990	1991	1992	1993	1994	1995	1996	1997
Total, T_n (millions)	57.41	57.57	57.73	57.89	58.05	58.21	58.37	58.53
Ratio T_n/T_{n-1}		1.003	1.003	1.003	1.003	1.003	1.003	1.003

From Table 3.5 it can be seen clearly that, according to the model, the population is increasing geometrically, with the ratio between successive years equal to 1.003 (to three decimal places).

This means that the annual proportionate growth rate is 0.003.

Apart from this general trend upwards, is there any other information to be gleaned from Table 3.4? To answer this, it is best to eliminate the general upward trend in each subpopulation, by considering the proportion of each in the total population. So we divide both subpopulation figures by the total population for the corresponding year. The results are tabulated in Table 3.6.

Table 3.6 Predictions of the subpopulation proportions from the model

	1990	1991	1992	1993	1994	1995	1996	1997
Proportion of juveniles, J_n/T_n	0.190	0.191	0.191	0.192	0.192	0.193	0.193	0.193
Proportion of adults, A_n/T_n	0.810	0.809	0.809	0.808	0.808	0.807	0.807	0.807

Looking at Table 3.6, you can see that the proportion of juveniles is gradually increasing. In Section 4, by plotting a graph on a computer, you will see that the rate of increase is slowing down. Also, by using a computer to calculate the predictions for the first fifty years, you will show that the proportion of juveniles approaches the limit 0.197, so the model predicts that eventually 19.7% of the population will be juveniles. The proportion of adults approaches the limit 0.803, so the model predicts that 80.3% of the population will be adults. So the answer to the question posed on page 29 is that the model predicts a stable situation where a constant proportion of juveniles enters the workforce each year, namely $19.7/15 = 1.3$ per cent of the population.

Note that the two proportions add to 1, because

$$\frac{J_n}{T_n} + \frac{A_n}{T_n} = \frac{T_n}{T_n} = 1.$$

So the proportions of juveniles and adults eventually settle down into constant proportions. You have already seen something analogous to this at the end of the video sequence. The last experiment involved cycling water repeatedly through a two-input, two-output network. It was found that the proportions quickly settled down.

Evaluate

To evaluate a model, more data are needed. The *Demographic Yearbook* (1997 edition) gives 11.36 million juveniles and 47.44 million adults in 1996. Comparing these figures with the values of $J_6 = 11.27$ and $A_6 = 47.10$ given in Table 3.4, the agreement is pretty good but not perfect, which suggests that the model should be refined. You will return to the more detailed three-subpopulation model in the next section, and will use it to answer the question 'Who will support us in our old age?' that was posed at the beginning of this section.

3.2 Closed-form solution

The matrix model of the subpopulations given by equation (3.2) has a closed-form solution. We derive this here.

Equation (3.2) can be expressed more succinctly as

$$\mathbf{p}_{n+1} = \mathbf{M}\mathbf{p}_n \quad (n = 0, 1, 2, \ldots), \tag{3.3}$$

where $\mathbf{M}$ is the matrix $\begin{pmatrix} 0.9326 & 0.0172 \\ 0.0666 & 0.9864 \end{pmatrix}$ and $\mathbf{p}_n$ is the vector $\begin{pmatrix} J_n \\ A_n \end{pmatrix}$.

Now consider how successive estimates of the subpopulations are obtained.

The population estimate for year 1 is obtained directly from the initial data, as $\mathbf{p}_1 = \mathbf{M}\mathbf{p}_0$. The population estimate for year 2 is obtained from the population estimate for year 1, as $\mathbf{p}_2 = \mathbf{M}\mathbf{p}_1$. Substituting for $\mathbf{p}_1$ from the previous equation gives the population estimate for year 2 in terms of the initial data: $\mathbf{p}_2 = \mathbf{M}(\mathbf{M}\mathbf{p}_0) = \mathbf{M}^2\mathbf{p}_0$. Similarly, the estimate for year 3 can be obtained from the initial data as $\mathbf{p}_3 = \mathbf{M}\mathbf{p}_2 = \mathbf{M}(\mathbf{M}^2\mathbf{p}_0) = \mathbf{M}^3\mathbf{p}_0$. Continuing in this way, we obtain the estimate for the population in year n in terms of the initial data, as

$$\mathbf{p}_n = \mathbf{M}^n\mathbf{p}_0. \tag{3.4}$$

This is the closed-form solution for the population problem.

You saw this approach for obtaining a closed-form solution in Chapter A1. There you saw that the recurrence equation

$$x_{n+1} = kx_n,$$

where x_n and k are real numbers, has the closed form

$$x_n = k^n x_0.$$

This was put to use in Chapter B1, in obtaining a closed-form solution for the exponential model of population change.

As it stands, the closed-form solution (3.4) is no improvement on the recurrence system (3.3) as in fact it takes more arithmetic operations to compute the former than the latter. The closed-form solution is useful, however, because there is a quick way of calculating powers of a matrix. This quick method is beyond the scope of this course, but is covered in second-level courses.

Summary of Section 3

This section has introduced a matrix model for exploring the ways in which two interdependent subpopulations behave, in order to gain some insight into the long-term behaviour of the structure of a population.

Exercise for Section 3

Exercise 3.1

The aim of this question is to develop a model of the population of a fictitious developing country. Table 3.7 gives the birth and death rates for two subpopulations of that country.

Table 3.7 Population of a fictitious developing country in January 2000: juveniles and adults

	Juveniles (aged < 15)	Adults (aged ≥ 15)
Subpopulation (millions)	47.85	614.37
Birth rate	0.0000	0.0520
Death rate	0.0139	0.0231

(a) Draw a network diagram similar to Figure 3.2, and label the pipes using the interpretations given in Table 3.3 and the data given in Table 3.7.

(b) Write down the corresponding matrix model for the population.

4 Computing with matrices

In this section you will be using Mathcad to create, add and multiply matrices. This work will increase your familiarity with the way matrices behave; also, it will help you to distinguish between circumstances when it is worth using Mathcad to perform matrix calculations, and when it would be more efficient to carry out the calculations on paper.

You will also investigate further the two-subpopulation model studied in Section 3, and solve the three-subpopulation model set up there. This model will enable you to address the question of whether in the long term there will be enough workers to support the ageing population.

The three-subpopulation model is encapsulated in the following network diagram and table, where at the start of year n, J_n is the number of juveniles, W_n is the number of workers, and E_n is the number of elderly.

Note that there are no pipes from J_n to E_{n+1} and E_n to W_{n+1}: juveniles cannot become elderly in one year, and the elderly cannot become younger, to become workers. There is also no link from E_n to J_{n+1}, on the assumption of a zero birth rate for the elderly.

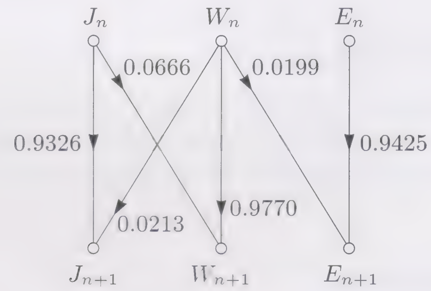


Figure 4.1 The interrelationships between the subpopulations in successive years

Table 4.1 Interpretation of the pipes in Figure 4.1

From	To	Interpretation (for year n)
J_n	J_{n+1}	Proportion of surviving juveniles who do not reach 15
W_n	W_{n+1}	Proportion of surviving workers who do not reach 65
E_n	E_{n+1}	Proportion of elderly who do not die
W_n	J_{n+1}	Proportionate birth rate for workers
J_n	W_{n+1}	Proportion of surviving juveniles who reach 15
W_n	E_{n+1}	Proportion of surviving workers who reach 65

The pipe labels in Figure 4.1 are calculated by using the interpretations in Table 4.1, as in Section 3 for the two-subpopulation model.

Refer to Computer Book B for the work in this section.

Summary of Section 4

This section provided practice in using Mathcad to create, add and multiply matrices. The population problems described in the previous section were studied in more detail.

5 Simultaneous linear equations and matrices

In this section we explore the close connection between simultaneous linear equations and matrices. In particular, you will see that simultaneous linear equations can be solved by using only matrix methods. We start the section by exploring a little more of the theory of matrices, and then apply it to solve simultaneous linear equations. In order to keep the arithmetic to a minimum, this section will be concerned mostly with 2×2 matrices.

5.1 Identity matrices

An *identity matrix* is a special type of matrix: multiplying a matrix by an identity matrix leaves the matrix unchanged, if the product exists. Do the following activity to see how this works.

Activity 5.1 A case of identity

Let $\mathbf{A} = \begin{pmatrix} 3 & 4 \\ -2 & 1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 2 & 7 & -3 \\ 8 & 1 & 2 \end{pmatrix}$ and $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Evaluate each of the following matrix products, where possible:

$$\mathbf{AI}, \quad \mathbf{IA}, \quad \mathbf{BI}, \quad \mathbf{IB}.$$

Solutions are given on page 51.

Let $\mathbf{I}$ be the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and let $\mathbf{A}$ be any matrix for which the product $\mathbf{AI}$ can be formed. (Thus $\mathbf{A}$ must have 2 columns.) Then it is always true that

$$\mathbf{AI} = \mathbf{A}.$$

What is more, if $\mathbf{B}$ is any matrix for which the product $\mathbf{IB}$ can be formed (that is, $\mathbf{B}$ is any matrix with 2 rows), then

$$\mathbf{IB} = \mathbf{B}.$$

This matrix $\mathbf{I}$ is called the 2×2 **identity matrix**.

Note that if $\mathbf{A}$ is a 2×2 matrix, then the products $\mathbf{AI}$ and $\mathbf{IA}$ can both be formed since both conditions above concerning rows and columns are met. So, for a 2×2 matrix $\mathbf{A}$, we have

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}.$$

There is only one 2×2 identity matrix (the matrix $\mathbf{I}$ used above), but there are also identity matrices of other sizes; for example,

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is the 3×3 identity matrix. This also has the properties $\mathbf{AI} = \mathbf{A}$ and $\mathbf{IB} = \mathbf{B}$, for suitable matrices $\mathbf{A}$ and $\mathbf{B}$ (see Activity 5.2 for a verification of this assertion).

The identity matrices of sizes 2×2 and 3×3 are sometimes written as $\mathbf{I}_2$ and $\mathbf{I}_3$, respectively.

$\mathbf{A}$ must have 3 columns, and $\mathbf{B}$ must have 3 rows.

Each identity matrix is a square matrix with ones down the **leading** (or **main**) diagonal (that is, the diagonal of the matrix which starts at the top left and ends at the bottom right) and every other element zero.

Activity 5.2 Identity matrices

Suppose that $\mathbf{I}$ is the 3×3 identity matrix and $\mathbf{b}$ is a vector with 3 components.

(a) According to the assertion in the text above, what is the matrix product $\mathbf{Ib}$?

(b) Now write $\mathbf{b}$ as

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

and calculate $\mathbf{Ib}$, thus verifying that the assertion is correct in this case.

Solutions are given on page 51.

5.2 Inverse of a matrix

This situation is analogous to the arithmetic of numbers. For example, dividing by 2 is equivalent to multiplying by $\frac{1}{2} = 2^{-1}$, which is the inverse of 2.

You may have realised by now that we have not divided a matrix by a matrix – there is no useful way of doing this. However, in some circumstances, multiplying by an *inverse* matrix suffices in situations where we would want to divide. This subsection describes what we mean by an inverse matrix and how to calculate the inverse of a 2×2 matrix. We start by looking at some special pairs of matrices.

Activity 5.3 Linked pairs

(a) Multiply the following pairs of matrices together. In each case, try reversing the order of multiplication ($\mathbf{BA}$ instead of $\mathbf{AB}$, for example), and see whether this affects the outcome.

(i) $\mathbf{A} = \begin{pmatrix} 8 & 5 \\ 3 & 2 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 2 & -5 \\ -3 & 8 \end{pmatrix}$

(ii) $\mathbf{C} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$

(iii) $\mathbf{P} = \begin{pmatrix} 3 & -9 \\ -1 & 4 \end{pmatrix}$, $\mathbf{Q} = \begin{pmatrix} 4 & 9 \\ 1 & 3 \end{pmatrix}$

(b) Describe the outcome of your calculations in part (a).

Solutions are given on page 51.

Suppose that $\mathbf{A}$ and $\mathbf{B}$ are any two 2×2 matrices with the property that

$$\mathbf{AB} = \mathbf{I} = \mathbf{BA},$$

where $\mathbf{I}$ is the 2×2 identity matrix (such as the matrices $\mathbf{A}$ and $\mathbf{B}$ in Activity 5.3). Then we say that $\mathbf{B}$ is the **inverse** of $\mathbf{A}$. We write the inverse of $\mathbf{A}$ as $\mathbf{A}^{-1}$, so we have

$$\mathbf{AA}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A}.$$

This equation shows that if $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$, then $\mathbf{A}$ must be the inverse of $\mathbf{A}^{-1}$. For the particular matrices in Activity 5.3, you showed that $\mathbf{A} = \mathbf{B}^{-1}$ and $\mathbf{B} = \mathbf{A}^{-1}$, and that $\mathbf{C} = \mathbf{D}^{-1}$ and $\mathbf{D} = \mathbf{C}^{-1}$. The matrices $\mathbf{P}$ and $\mathbf{Q}$ are not the inverses of each other, however, since their product is *not* an identity matrix.

Even though the matrices $\mathbf{P}$ and $\mathbf{Q}$ are not inverses of each other, they are special in that the product of $\mathbf{P}$ and $\mathbf{Q}$ is a scalar multiple of the identity matrix, namely $3\mathbf{I}$. Since $\mathbf{PQ} = 3\mathbf{I}$,

$\mathbf{PQ} = \mathbf{QP} = 3\mathbf{I}$

$$\tfrac{1}{3}(\mathbf{PQ}) = \tfrac{1}{3}(3\mathbf{I}) = \mathbf{I}.$$

Using the result that $\mathbf{A}(k\mathbf{B}) = k(\mathbf{AB})$, given in equation (2.2), we have

$$\mathbf{P}(\tfrac{1}{3}\mathbf{Q}) = \tfrac{1}{3}(\mathbf{PQ}) = \mathbf{I}.$$

Similarly, $(\tfrac{1}{3}\mathbf{Q})\mathbf{P} = \mathbf{I}$. So the inverse of $\mathbf{P}$ is obtained by scalar multiplying $\mathbf{Q}$ by $\tfrac{1}{3}$. This 'trick' of scalar multiplying matrices appropriately can be used to find a general formula for the inverse of a 2×2 matrix. The next activity prepares the ground.

Similarly, $\tfrac{1}{3}\mathbf{P}$ is the inverse of $\mathbf{Q}$.

Activity 5.4 Deriving a formula for the inverse of a 2×2 matrix

A general 2×2 matrix can be written in the form $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c and d are real numbers. Find the product

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

A solution is given on page 52.

Comment

Element notation has not been used in this example, since it would be cumbersome to do so.

The product

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

in Activity 5.4 yields a scalar multiple of the identity matrix, namely $(ad - bc)\mathbf{I}$, so scalar multiplying the second matrix above by $1/(ad - bc)$ gives a formula for the inverse of a general 2×2 matrix.

Inverse of a 2×2 matrix

The inverse of the general 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is given by

$$\frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad \text{provided that } ad - bc \neq 0.$$

The following example shows how this formula can be used to compute the inverse of a matrix.

Example 5.1 Finding an inverse

Find the inverse of the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 7 \\ 3 & 6 \end{pmatrix},$$

and check that your answer *is* the inverse.

Solution

By comparing the matrix $\mathbf{A}$ given in the question with the general 2×2 matrix, we see that $a = 4$, $b = 7$, $c = 3$ and $d = 6$. So

$$ad - bc = (4 \times 6) - (7 \times 3) = 24 - 21 = 3,$$

and the inverse of $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \tfrac{1}{3}\begin{pmatrix} 6 & -7 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & -\frac{7}{3} \\ -1 & \frac{4}{3} \end{pmatrix}.$$

To check that this *is* the required inverse matrix, we multiply it by $\mathbf{A}$:

$$\begin{aligned} \begin{pmatrix} 2 & -\frac{7}{3} \\ -1 & \frac{4}{3} \end{pmatrix}\begin{pmatrix} 4 & 7 \\ 3 & 6 \end{pmatrix} &= \begin{pmatrix} 2 \times 4 + (-\frac{7}{3}) \times 3 & 2 \times 7 + (-\frac{7}{3}) \times 6 \\ (-1) \times 4 + \frac{4}{3} \times 3 & (-1) \times 7 + \frac{4}{3} \times 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

This product is the identity matrix, so we have checked that $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$.

Square matrices of size bigger than 2×2 also have determinants, but they are not discussed in this course.

It is useful to be able to refer to the quantity $ad - bc$: it is called the **determinant** of the 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

and is often written $\det \mathbf{A}$. Thus

$$\det \mathbf{A} = ad - bc.$$

If the determinant of a matrix is zero, then the above formula for the inverse of the matrix fails (because division by zero is not defined). Not only does the formula fail but, in fact, a matrix with zero determinant does not have an inverse, as you will see geometrically in the next subsection. This role of the determinant is summarised in the following test.

Other terms in widespread use are *non-singular* instead of invertible and *singular* instead of non-invertible.

Determinant Test

If the determinant of a matrix $\mathbf{A}$ is not zero, then $\mathbf{A}$ has an inverse and we say that $\mathbf{A}$ is **invertible**.

If the determinant of a matrix $\mathbf{A}$ is zero, then $\mathbf{A}$ does not have an inverse and we say that $\mathbf{A}$ is **non-invertible**.

For example, the matrix $\mathbf{B} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}$ has determinant

$$\det \mathbf{B} = 3 \times 8 - 4 \times 6 = 24 - 24 = 0.$$

Thus $\mathbf{B}$ is non-invertible; that is, $\mathbf{B}^{-1}$ does not exist.

The next activity offers practice in identifying 2×2 matrices that have an inverse and in finding inverse matrices.

Activity 5.5 Finding inverses

Use the Determinant Test to classify each of the following matrices as invertible or non-invertible. For each of the invertible matrices, write down its inverse.

(a) $\mathbf{A} = \begin{pmatrix} 13 & 5 \\ 5 & 2 \end{pmatrix}$ (b) $\mathbf{B} = \begin{pmatrix} 3 & 0 \\ -5 & 2 \end{pmatrix}$

(c) $\mathbf{C} = \begin{pmatrix} -3 & \frac{1}{5} \\ 5 & -\frac{1}{3} \end{pmatrix}$ (d) $\mathbf{D} = \begin{pmatrix} 1.5 & 2.5 \\ 0.5 & 1.5 \end{pmatrix}$

Solutions are given on page 52.

5.3 Simultaneous linear equations

In Chapter A0, two methods (substitution and elimination) for solving pairs of simultaneous linear equations were described. Here a third method, which uses matrices, is described.

An example of a pair of simultaneous linear equations in the two variables x and y is

$$\begin{aligned} 0.4x + 0.2y &= 4, \\ 0.6x + 0.8y &= 11. \end{aligned} \tag{5.1}$$

Each of these equations represents a straight line in the (x, y)-plane. You may not recognise the above equations as representing straight lines, but they can both be rearranged into the more familiar form $y = mx + c$. This can be done by subtracting the x-term from both sides and then dividing by the coefficient of y. So the first equation above becomes

The representation of a straight line by an equation of the form $y = mx + c$ was discussed in Chapter A2, Subsection 1.1.

$$y = \frac{4 - 0.4x}{0.2} = -2x + 20,$$

and the second equation becomes

$$y = \frac{11 - 0.6x}{0.8} = -\tfrac{3}{4}x + \tfrac{55}{4}.$$

These two straight lines are shown in Figure 5.1, overleaf.

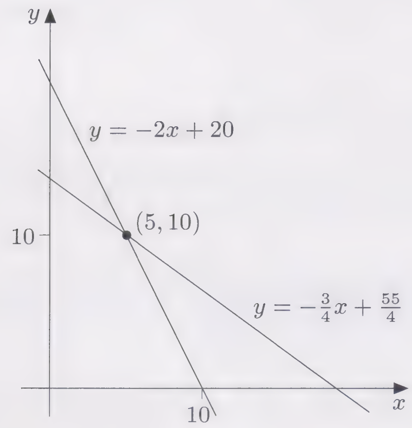


Figure 5.1 The two straight lines which correspond to equations (5.1)

Equations (5.1) can be thought of as one equation connecting two vectors, rather than two equations connecting numbers, by writing each equation as one component of a vector, as follows:

$$\begin{pmatrix} 0.4x + 0.2y \\ 0.6x + 0.8y \end{pmatrix} = \begin{pmatrix} 4 \\ 11 \end{pmatrix}.$$

Re-writing the left-hand side of this equation using a matrix–vector product gives the matrix form of equations (5.1):

$$\begin{pmatrix} 0.4 & 0.2 \\ 0.6 & 0.8 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 11 \end{pmatrix}. \qquad (5.2)$$

We shall need to refer frequently to the 2×2 matrix that appears in the above equation, so we call it the **coefficient matrix** of the simultaneous linear equations (5.1). Note that the elements of the coefficient matrix are the same numbers arranged in the same order as the coefficients on the left-hand sides of equations (5.1). This will always be the case for the coefficient matrix of any pair of simultaneous linear equations in which the variables appearing on the left-hand sides are written down in the same order in each equation.

The following activity is designed to give you practice in writing down the matrix form of a pair of simultaneous linear equations.

Activity 5.6 Matrix form of simultaneous linear equations

Rewrite each of the following pairs of equations in matrix form.

(a) $2x + 3y = 3$
$\quad x + 4y = -1$

(b) $2x \qquad = 6$
$\quad 3x + 6y = 15$

(c) $\frac{3}{5}x - \frac{4}{5}y = 18$
$\quad \frac{4}{5}x + \frac{3}{5}y = -1$

Solutions are given on page 52.

We now return to equation (5.2), the matrix form of equations (5.1), and proceed to solve the equations. Let $\mathbf{A}$ be the coefficient matrix, that is,

$$\mathbf{A} = \begin{pmatrix} 0.4 & 0.2 \\ 0.6 & 0.8 \end{pmatrix}.$$

The determinant of $\mathbf{A}$ is $\det \mathbf{A} = 0.2$, so the inverse of $\mathbf{A}$ is

$\det \mathbf{A} = 0.4 \times 0.8 - 0.2 \times 0.6$
$= 0.32 - 0.12 = 0.2.$

$$\mathbf{A}^{-1} = \frac{1}{0.2}\begin{pmatrix} 0.8 & -0.2 \\ -0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix}.$$

Now multiply both sides of equation (5.2) by $\mathbf{A}^{-1}$, to obtain

$$\mathbf{A}^{-1}\mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A}^{-1}\begin{pmatrix} 4 \\ 11 \end{pmatrix}. \qquad (5.3)$$

By definition, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, since $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$, and $\mathbf{I}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$, since multiplying a matrix by an identity matrix leaves that matrix unchanged. Hence, on simplifying the left-hand side and evaluating the right-hand side of equation (5.3), we obtain

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 11 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix}.$$

In other words,

$$x = 5 \quad \text{and} \quad y = 10.$$

This solution can be checked by substituting the values for x and y into equations (5.1).

The example above illustrates a general method for solving simultaneous linear equations which can be summarised as follows.

Solving a pair of simultaneous linear equations using matrices

First write the simultaneous linear equations in matrix form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is the coefficient matrix, $\mathbf{x}$ is the corresponding vector of variables, and $\mathbf{b}$ is the vector with components equal to the corresponding right-hand sides of the equations.

If the matrix $\mathbf{A}$ is invertible, then the solution is given by

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

The next activity provides practice in using this method to solve simultaneous linear equations.

Activity 5.7 Solving simultaneous linear equations

Use matrices to solve the following pairs of simultaneous linear equations. (*Hint*: You have already written each pair of equations in matrix form in Activity 5.6.)

(a) $2x + 3y = 3$
$x + 4y = -1$

(b) $2x = 6$
$3x + 6y = 15$

(c) $\frac{3}{5}x - \frac{4}{5}y = 18$
$\frac{4}{5}x + \frac{3}{5}y = -1$

Solutions are given on page 52.

Each equation of a pair of simultaneous linear equations represents a straight line. So the solution to a pair of simultaneous linear equations must correspond to a point that is simultaneously on both lines, that is, the point of intersection. This is illustrated in Figure 5.2, which shows the lines which represent the pairs of equations that you solved in Activity 5.7.

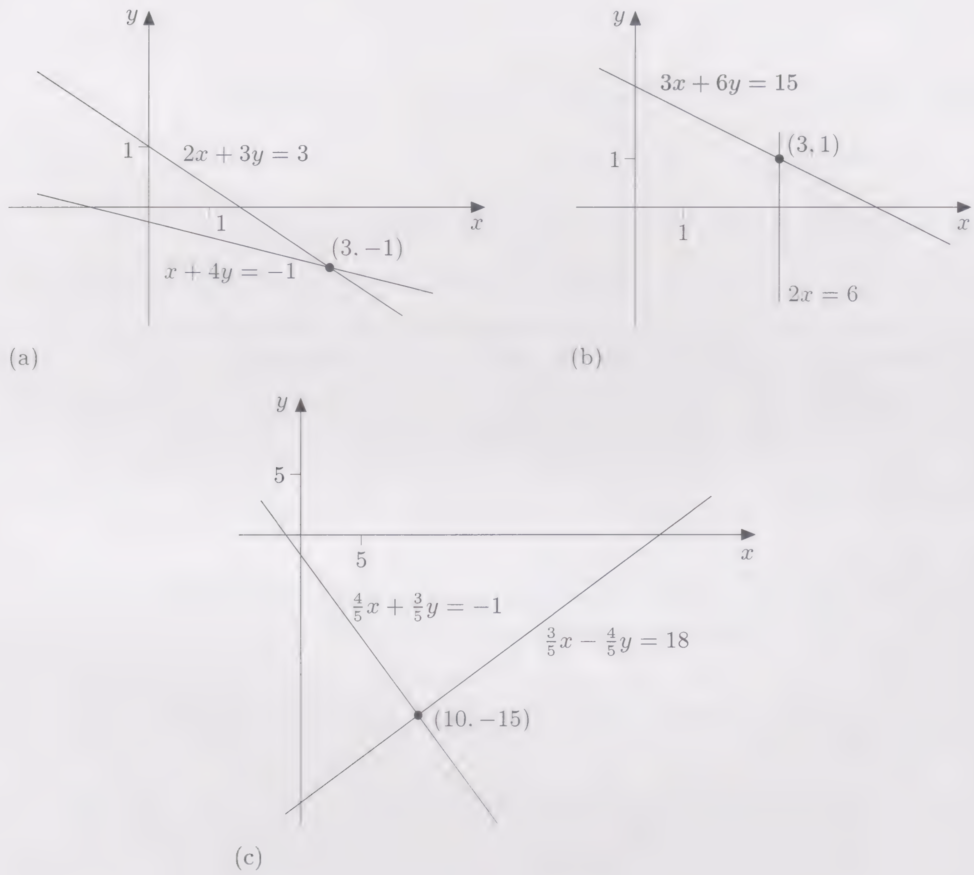


Figure 5.2 Points of intersection

An equation in n variables is *linear* if the largest power of any variable is 1, and no products of variables are present. For example,

$$x + 2y + z = 3$$

is linear, but neither

$$x + y^2 = 2$$

nor

$$x + yz = 1$$

is linear.

The matrix method of solution can be used for larger systems of linear equations. In general, a system of n simultaneous linear equations in n unknowns can be written as a single equation in matrix form in which a known n-component vector $\mathbf{b}$ is equal to an unknown vector $\mathbf{x}$ multiplied by an $n \times n$ coefficient matrix $\mathbf{A}$:

$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

If $\mathbf{A}$ is invertible, then this equation may be solved by multiplying both sides by the inverse of $\mathbf{A}$, giving $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. The drawback of this method is that finding an inverse is a lengthy calculation for any matrix larger than the 2×2 matrices considered here.

Exceptional cases

Most pairs of simultaneous linear equations are like the ones considered above and have a *unique* solution. However, this is not always the case, and it is the exceptions that concern us next.

We know that some 2×2 matrices do not have inverses. In the next activity you are asked to consider what this means in the context of simultaneous linear equations.

Activity 5.8 Non-invertible coefficient matrices

Rewrite the following pairs of equations in matrix form. Show that in each case the coefficient matrix does not have an inverse. Try to solve the equations using the elimination method: what happens?

(a) $2x - 3y = 5$
$-4x + 6y = 7$

(b) $x + 2y = -6$
$-3x - 6y = 18$

Solutions are given on page 53.

You saw in the previous activity two pairs of equations for which the coefficient matrix did not have an inverse, so the matrix method of finding solutions would not work. You also saw that the elimination method provided no solutions. In fact, no method for finding a unique solution will work, because neither pair of equations has a unique solution. To demonstrate this visually, the straight lines that correspond to the equations have been plotted in Figure 5.3.

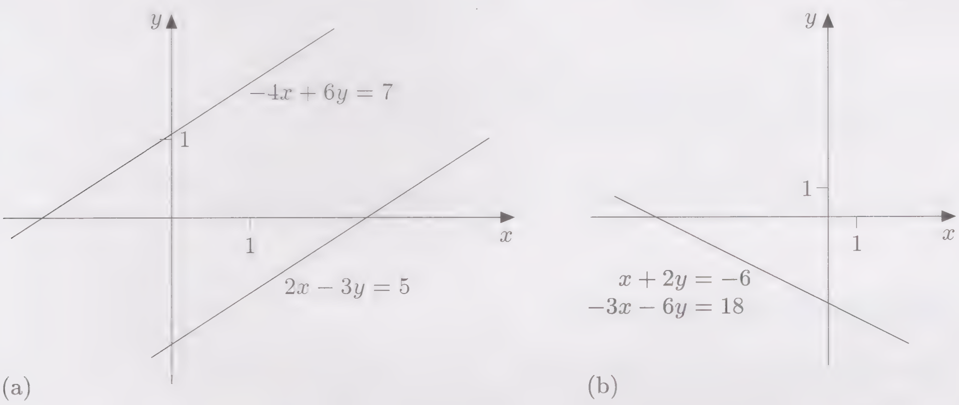


Figure 5.3 Exceptional cases

As you can see from Figure 5.3, there are two cases for which there is no unique solution. In Figure 5.3(a) the two lines are parallel, so they have no point of intersection. This shows that the equations have no solution. In Figure 5.3(b) both equations describe the same straight line, namely $y = -x/2 - 3$. So the 'two lines' intersect at every point along the line, and there are infinitely many solutions.

If the coefficient matrix is non-invertible, then one of the two possibilities described above occurs: either there is no solution or there are infinitely many solutions. In this course we do not elaborate on the methods of distinguishing between these two possibilities, beyond the graphical method used above.

A method for distinguishing between the two possibilities is described in second-level mathematics courses. The method is based on the 'elimination' method described in Chapter A0, and is called Gaussian elimination.

Final remarks

The discussion in this subsection has shown how to solve pairs of simultaneous linear equations by finding the inverse of the coefficient matrix. The methods described in Chapter A0 are preferred for hand calculations, because these methods involve fewer arithmetic operations than the matrix method. Note, however, that you may be specifically asked to use the matrix method for hand calculations in the assessment of this course, because it provides a good test of your understanding of matrices.

For computer calculations, a method similar to the 'elimination' hand calculation method (described in Chapter A0) is quicker than finding the inverse of the coefficient matrix.

So for what is the matrix form of the solution useful? The main use is in developing the theory of simultaneous linear equations. As the first payoff from this theory, we can say that simultaneous linear equations have a unique solution if and only if the coefficient matrix is invertible. Using the Determinant Test for invertibility then gives a quick test for the existence of a unique solution (or, equivalently, for a pair of lines having a unique point of intersection). Try this now in the following activity.

Activity 5.9 Determinant Test for a unique solution

Which of the following pairs of simultaneous linear equations has a unique solution? (You are not asked to find the solution.)

(a) $\begin{aligned} 5x - 3y &= 5 \\ -x + 4y &= 7 \end{aligned}$ (b) $\begin{aligned} 3x - y &= -6 \\ -9x + 3y &= 18 \end{aligned}$ (c) $\begin{aligned} 6x - 2y &= 1 \\ 3x - 7y &= -1 \end{aligned}$

Solutions are given on page 53.

Now would be a good time to listen to the *optional* band of Audio Tape 2, which describes some applications of matrices.

Listen to Audio Tape 2, Band 1, 'Applications of matrices'.

Summary of Section 5

This section has introduced:

- ◇ identity matrices;
- ◇ the Determinant Test for identifying invertible matrices;
- ◇ the formula for the inverse of an invertible 2×2 matrix;
- ◇ expressing a pair of simultaneous linear equations in matrix form as $\mathbf{Ax} = \mathbf{b}$;
- ◇ a method for solving a pair of simultaneous linear equations $\mathbf{Ax} = \mathbf{b}$, in the case where $\mathbf{A}$ is invertible.

Exercises for Section 5

Exercise 5.1

Find the inverse of each of the following matrices.

(a) $\mathbf{A} = \begin{pmatrix} -3 & -5 \\ 1 & 4 \end{pmatrix}$ (b) $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 5 & 5 \end{pmatrix}$ (c) $\mathbf{C} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}$

Exercise 5.2

Solve the following pairs of simultaneous linear equations using matrix methods.

(a) $\begin{aligned} -3x - 5y &= 1 \\ x + 4y &= 2 \end{aligned}$ (b) $\begin{aligned} 6x + 2y &= 4 \\ 3x - y &= 10 \end{aligned}$ (c) $\begin{aligned} 4x + 2y &= 2 \\ x + 3y &= -3 \end{aligned}$

Summary of Chapter B2

In this chapter you met the concept of matrices. You have used matrices to investigate a population case-study and also to solve pairs of simultaneous linear equations.

Learning outcomes

You have been working towards the following learning outcomes.

Terms to know and use

Network, vector, component of a vector, matrix, element of a matrix, row of a matrix, column of a matrix, size of a matrix, matrix addition, matrix multiplication, square matrix, powers of a matrix, scalar multiplication of matrices, identity matrix, inverse of a matrix, invertible matrix, non-invertible matrix, determinant of a matrix.

Symbols and notation to know and use

$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\mathbf{Av}$, $\mathbf{AB}$, $\mathbf{A}^n$, $\mathbf{A}+\mathbf{B}$,

$k\mathbf{A}$, $\mathbf{A}-\mathbf{B}$, $\mathbf{I}$, $\mathbf{A}^{-1}$, $\det \mathbf{A}$

Mathematical skills

- ◇ Represent networks by matrices.
- ◇ Describe a matrix in terms of the numbers of its rows and columns.
- ◇ Write down an example of an $m \times n$ matrix for given values of m and n.
- ◇ Evaluate matrix products.
- ◇ Evaluate matrix sums.
- ◇ Evaluate the square of a matrix.
- ◇ Evaluate the result of scalar multiplying a matrix.
- ◇ Use element notation to describe matrix elements in terms of their row and column positions.
- ◇ Write down conditions under which it is possible to form matrix products, sums and powers.
- ◇ Decide whether a 2×2 matrix has an inverse, and evaluate the inverse (if it exists).
- ◇ Write a pair of simultaneous linear equations in matrix form, and solve them using matrix methods where possible.

Mathcad skills

- ◇ Create matrices.
- ◇ Add and multiply matrices.
- ◇ Scalar multiply matrices and evaluate powers of matrices.
- ◇ Calculate a sequence of vectors using matrix–vector multiplication.

Solutions to Activities

Solution 1.1

(a) In Figure 1.1 the label on the pipe from node B to node U is 0.3. So 0.3 litres of water are output at node U.

Similarly, by reading the labels on the other pipes from node B, we have 0.1 litres of water output at node V, and 0.6 litres of water output at node W.

(b) First consider inputting 1 litre of water at node A. Then, as described in the text preceding the activity, 0.4 litres of water will be output at node U. Now add another 1 litre of water to node A. This will add another 0.4 litres of water to the 0.4 litres already output at node U. So, all together, $2 \times 0.4 = 0.8$ litres of water will be output at node U.

Similarly, the outputs at the other two nodes will double. So $2 \times 0.2 = 0.4$ litres of water will be output at node V, and $2 \times 0.4 = 0.8$ litres of water will be output at node W.

Solution 1.2

(a) Adding together the contribution from each input gives:

$0.4 + (2 \times 0.3) = 1$ litre of water output at node U;

$0.2 + (2 \times 0.1) = 0.4$ litres of water output at node V;

$0.4 + (2 \times 0.6) = 1.6$ litres of water output at node W.

(b) First consider the amount of water reaching node U from node A. 1 litre input gives 0.4 litres output, so x litres input gives $0.4x$ litres output. Similarly, the amount of water reaching node U from node B is $0.3y$ litres. Adding together the contributions from the two input nodes gives $0.4x + 0.3y$ litres of water output at node U.

Similarly, we add together the contributions from the two input nodes to obtain:

$0.2x + 0.1y$ litres of water output at node V;

$0.4x + 0.6y$ litres of water output at node W.

Solution 1.3

(a) Inputting 1 litre of water at node U gives 0.2 litres of water output at node C and 0.8 litres of water output at node D.

Inputting 2 litres of water at node V gives 2×0.6 litres of water output at node C and 2×0.4 litres of water output at node D.

So, in total, we have $0.2 + 2 \times 0.6 = 1.4$ litres of water output at node C, and $0.8 + 2 \times 0.4 = 1.6$ litres of water output at node D.

(b) Using a method similar to that of part (a) gives $0.2x + 0.6y + 0.3z$ litres of water output at node C and $0.8x + 0.4y + 0.7z$ litres of water output at node D.

Solution 1.4

(a) There are three routes through the network that start at node A and end at node D, namely the routes through nodes U, V and W. The method of calculating the output at node D is to add together the contributions from each of the possible routes.

First consider the route that goes via node U. From Figure 1.3, 0.4 litres of water flow from node A to node U. Also from the figure, if 1 litre were input at node U, then 0.8 litres would be output at node D. So if 0.4 litres are input at node U, then $0.4 \times 0.8 = 0.32$ litres are output at node D.

Similarly, the amount of water flowing from node A to node D via node V is $0.2 \times 0.4 = 0.08$ litres, and the amount of water flowing from node A to node D via node W is $0.4 \times 0.7 = 0.28$ litres.

So, in total, the amount of water output at node D is

$$0.32 + 0.08 + 0.28 = 0.68 \text{ litres.}$$

(b) As in part (a), the amounts of water output can be found by calculating the flow through each of the possible routes. 1 litre of water input at node B gives:

$(0.3 \times 0.2) + (0.1 \times 0.6) + (0.6 \times 0.3) = 0.3$ litres output at node C;

$(0.3 \times 0.8) + (0.1 \times 0.4) + (0.6 \times 0.7) = 0.7$ litres output at node D.

Solution 1.5

(a) The answer can be read off from Figure 1.4: 0.3 litres of water are output at node C, and 0.7 litres of water are output at node D. (This is much easier than calculating the outputs using the combined network, as was done in Activity 1.4(b).)

(b) From Figure 1.4, we find the outputs from each input and then add together the contributions. This gives $0.32x + 0.3y$ litres of water output at node C, and $0.68x + 0.7y$ litres of water output at node D.

Solution 1.6

(a) The first component of the vector is the number of litres of water input at node A, which in this case is 2. The second component of the vector is the number of litres of water input at node B, which in this case is 3. So the required vector is

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

(b) The first component of the vector is 2.

(c) The three-component vector given in the question represents outputs of 0.3 litres of water at node U, 0.1 litres of water at node V, and 0.6 litres of water at node W.

Solution 1.7

(a) The vector which represents an input of 1 litre of water at node V and no other inputs is

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

(Note that the input vector has three components because there are three input nodes in the network diagram.)

(b) The input vector has three components. So, according to the convention, the matrix representing the network must have three columns (one for each input node).

(c) The output vector has two components, since there are two output nodes. So, according to the convention, the matrix representing the network must have two rows.

(d) By parts (b) and (c), the matrix representing the network in Figure 1.6 must have two rows and three columns. The number that labels the pipe from the first input node (U) to the first output node (C) is 0.2. This is the number that should go in the first row and first column. The number that labels the pipe from the second input node (V) to the first output node (C) is 0.6, and this should be put in the first row and second column.

Proceeding similarly for all the inputs to and outputs from the network gives the following matrix to represent the network:

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix}.$$

Solution 1.8

(a) The required input vector is $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

(b) Start by writing down the matrix representing the network and the input vector:

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Now use the rule which was described in the text preceding the activity: for each row of the matrix, add together the results of multiplying each element in the row with the corresponding element in the input vector. In this case, this procedure gives

$$\begin{pmatrix} 0.2 \times 1 + 0.6 \times 0 + 0.3 \times 0 \\ 0.8 \times 1 + 0.4 \times 0 + 0.7 \times 0 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}.$$

This is the output vector, which represents an output of 0.2 litres of water at node C and 0.8 litres of water at node D.

Solution 1.9

(a)
$$\begin{aligned} \begin{pmatrix} 0.33 & 0.25 \\ 0.67 & 0.75 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 0.33 + 0.25 \\ 0.67 + 0.75 \end{pmatrix} \\ &= \begin{pmatrix} 0.58 \\ 1.42 \end{pmatrix} \end{aligned}$$

(b)
$$\begin{aligned} \begin{pmatrix} 0.33 & 0.25 \\ 0.67 & 0.75 \end{pmatrix} \begin{pmatrix} 10 \\ 14 \end{pmatrix} &= \begin{pmatrix} 0.33 \times 10 + 0.25 \times 14 \\ 0.67 \times 10 + 0.75 \times 14 \end{pmatrix} \\ &= \begin{pmatrix} 3.3 + 3.5 \\ 6.7 + 10.5 \end{pmatrix} \\ &= \begin{pmatrix} 6.8 \\ 17.2 \end{pmatrix} \end{aligned}$$

(c)
$$\begin{aligned} \begin{pmatrix} 0.33 & 0.25 \\ 0.67 & 0.75 \end{pmatrix} \begin{pmatrix} 18 \\ 6 \end{pmatrix} &= \begin{pmatrix} 0.33 \times 18 + 0.25 \times 6 \\ 0.67 \times 18 + 0.75 \times 6 \end{pmatrix} \\ &= \begin{pmatrix} 5.94 + 1.5 \\ 12.06 + 4.5 \end{pmatrix} \\ &= \begin{pmatrix} 7.44 \\ 16.56 \end{pmatrix} \end{aligned}$$

Solution 1.10

(a) (i) The first task is to work out which nodes are involved in the top right element of the product matrix. The top row corresponds to the first output, so we are looking for routes which end at node C. The right-hand column corresponds to the second input, so we are looking for routes which start at node B. The network diagram with routes from node B to node C highlighted is given in Figure S.1.

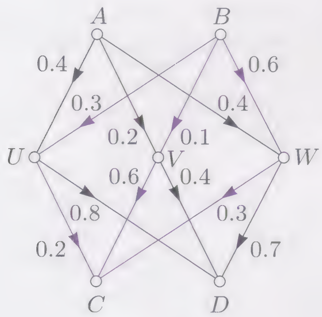


Figure S.1

(ii) The relevant elements in the two matrices are the top row of the first matrix and the right-hand column of the second matrix, as highlighted in Figure S.2.

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix}$$

Figure S.2

(iii) Adding the products of corresponding elements gives

$$0.2 \times 0.3 + 0.6 \times 0.1 + 0.3 \times 0.6 = 0.3.$$

So the top right element of the product matrix is 0.3, which agrees with the value shown in Figure 1.9.

(b) (i) This time, the bottom right element of the product is required; this corresponds to routes between the second input node (B) and the second output node (D). The network diagram with routes from node B to node D highlighted is given in Figure S.3.

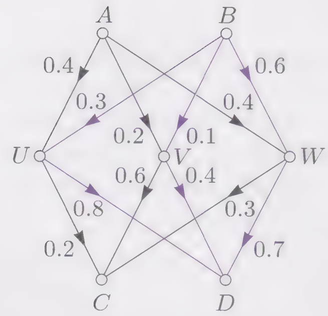


Figure S.3

(ii) The relevant elements in the two matrices are the bottom row of the first matrix and the right-hand column of the second matrix, as highlighted in Figure S.4.

$$\begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.8 & 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \\ 0.4 & 0.6 \end{pmatrix}$$

Figure S.4

(iii) Adding the products of corresponding elements gives

$$0.8 \times 0.3 + 0.4 \times 0.1 + 0.7 \times 0.6 = 0.7.$$

So the bottom right element of the product matrix is 0.7, which agrees with the value shown in Figure 1.9.

Solution 2.1

In order to use the method described in the text, the sizes of the matrices are written underneath the matrices, and the middle two numbers are boxed.

(a)
$$\begin{pmatrix} \frac{1}{2} & 9 \\ 7 & \frac{1}{3} \\ -6 & 8 \end{pmatrix} \begin{pmatrix} -2 & 3 & 1 \\ 4 & 0 & 5 \end{pmatrix}$$
$$3 \times \boxed{2 \quad 2} \times 3$$

The two numbers in the box are equal, so the matrices can be multiplied. Ignoring the numbers in the box gives the size of the product matrix as 3×3. (So the product matrix will have 3 rows and 3 columns.)

(b)
$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 9 \\ 8 \\ 7 \end{pmatrix}$$
$$1 \times \boxed{3 \quad 3} \times 1$$

The two numbers in the box are equal, so the matrices can be multiplied. Ignoring the numbers in the box gives the size of the product matrix as 1×1. (So the product matrix will have 1 row and 1 column. By convention, we usually omit the parentheses around a 1×1 matrix and treat it as a number.)

(c) $$\begin{pmatrix} \frac{1}{2} & 9 \\ 7 & \frac{1}{3} \\ -6 & 8 \end{pmatrix}\begin{pmatrix} 0 & 10 \\ 4 & \frac{1}{2} \\ -2 & 7 \end{pmatrix}$$

$$3 \times \boxed{2 \quad 3} \times 2$$

The two numbers in the box do not match, so these two matrices cannot be multiplied together.

(d) $$\begin{pmatrix} 9 \\ 8 \\ 7 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$$

$$3 \times \boxed{1 \quad 1} \times 3$$

The two numbers in the box are equal, so the matrices can be multiplied. Ignoring the numbers in the box gives the size of the product matrix as 3×3. (So the product matrix will have 3 rows and 3 columns.)

(e) $$\begin{pmatrix} 1 & 2 & 3 \\ 7 & 6 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 2 & 1 & 3 \end{pmatrix}$$

$$2 \times \boxed{3 \quad 3} \times 3$$

The two numbers in the box are equal, so the matrices can be multiplied. Ignoring the numbers in the box gives the size of the product matrix as 2×3. (So the product matrix will have 2 rows and 3 columns.)

Solution 2.2

(a) $$\mathbf{AB} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \\ 6 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 5 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 5 & 11 \\ 6 & 10 & 22 \\ 7 & 17 & 31 \end{pmatrix}$$

$$3 \times \boxed{2 \quad 2} \times 3 \quad \text{gives} \quad 3 \times 3$$

$$\mathbf{BA} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 4 & 2 \\ 6 & 1 \end{pmatrix} = \begin{pmatrix} 44 & 12 \\ 4 & 0 \end{pmatrix}$$

$$2 \times \boxed{3 \quad 3} \times 2 \quad \text{gives} \quad 2 \times 2$$

(b) $$\mathbf{AB} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 4 & 6 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 10 & 11 \\ 12 & 18 \end{pmatrix}$$

$$2 \times \boxed{2 \quad 2} \times 2 \quad \text{gives} \quad 2 \times 2$$

$$\mathbf{BA} = \begin{pmatrix} 4 & 6 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 26 & 4 \\ 1 & 2 \end{pmatrix}$$

$$2 \times \boxed{2 \quad 2} \times 2 \quad \text{gives} \quad 2 \times 2$$

(c) $$\mathbf{AB} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \\ 4 & 6 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 8 \\ 9 & 12 \\ 22 & 32 \end{pmatrix}$$

$$3 \times \boxed{2 \quad 2} \times 2 \quad \text{gives} \quad 3 \times 2$$

BA cannot be formed, since the number of columns (2) in **B** does not match the number of rows (3) in **A**. Writing out the sizes of the matrices gives

$$2 \times \boxed{2 \quad 3} \times 2,$$

which also confirms that the product **BA** cannot be formed, because the numbers in the box do not match.

Solution 2.3

(a) The matrix product **AB** can be formed only if the number of columns of **A** is equal to the number of rows of **B**. Since the number of columns (3) of **M** is not equal to the number of rows (2) of **M**, the product **MM** cannot be formed.

(b) **A** must have the same number of rows as columns in order for it to be possible to calculate its square.

Solution 2.4

(a) If $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 4 & 6 \\ 2 & -1 \end{pmatrix}$, then

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} = \begin{pmatrix} 6 & 7 \\ 5 & -1 \end{pmatrix}.$$

(b) If $\mathbf{A} = \begin{pmatrix} 2 & 0 & 4 \\ -1 & 6 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, then

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} = \begin{pmatrix} 3 & 0 & 5 \\ -1 & 7 & 1 \end{pmatrix}.$$

(c) If $\mathbf{A} = \begin{pmatrix} 4 & 0 & 5 \\ 6 & 1 & 0 \\ 2 & 3 & -1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 6 & 2 & 0 \\ 0 & -1 & 4 \\ 5 & 0 & 1 \end{pmatrix}$, then

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} = \begin{pmatrix} 10 & 2 & 5 \\ 6 & 0 & 4 \\ 7 & 3 & 0 \end{pmatrix}.$$

Solution 2.5

(a) (i) $$\begin{pmatrix} 0.83 \\ 0.17 \end{pmatrix} + \begin{pmatrix} 0.36 \\ 0.64 \end{pmatrix} = \begin{pmatrix} 1.19 \\ 0.81 \end{pmatrix}$$

(ii) $$\begin{pmatrix} -5 \\ 7 \\ -4 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} -8 \\ 8 \\ 0 \end{pmatrix}$$

(iii) $$\begin{pmatrix} 6 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 \\ 2 \end{pmatrix} + \begin{pmatrix} 9 \\ 3 \end{pmatrix} = \begin{pmatrix} 13 \\ 6 \end{pmatrix}$$

(b) (i) Any pair of vectors with different numbers of components will do. For example, $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 5 \end{pmatrix}$ cannot be added together, since the first vector has three components and the second has only two.

Solution 2.6

(a) (i) $3\begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 12 & 9 \\ 6 & 3 \end{pmatrix}$

(ii) $\frac{1}{2}\begin{pmatrix} 10 & -3 \\ 7 & 3 \end{pmatrix} = \begin{pmatrix} 5 & -\frac{3}{2} \\ \frac{7}{2} & \frac{3}{2} \end{pmatrix}$

(iii) $\frac{2}{3}\begin{pmatrix} 6 & x \\ 3 & y \end{pmatrix} = \begin{pmatrix} 4 & \frac{2}{3}x \\ 2 & \frac{2}{3}y \end{pmatrix}$

(b) (i) $\begin{pmatrix} 4.5 & 3 \\ 2 & 1.5 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 9 & 6 \\ 4 & 3 \end{pmatrix}$

(ii) $\begin{pmatrix} \frac{3}{2} \\ \frac{2}{3} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 9 \\ 4 \end{pmatrix}$

(iii) $\begin{pmatrix} 2x & 3x \\ 0 & 5x \end{pmatrix} = x\begin{pmatrix} 2 & 3 \\ 0 & 5 \end{pmatrix}$

In each of the answers in this part, the scalar multiple is not unique. For example, in part (i),

$$\begin{pmatrix} 4.5 & 3 \\ 2 & 1.5 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 18 & 12 \\ 8 & 6 \end{pmatrix}.$$

(c) Since $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$, $2\mathbf{B} = \begin{pmatrix} 2 & 0 \\ 4 & 2 \end{pmatrix}$. Hence

$$\mathbf{A}(2\mathbf{B}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ 22 & 8 \end{pmatrix}.$$

Also,

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 11 & 4 \end{pmatrix},$$

hence

$$2(\mathbf{AB}) = \begin{pmatrix} 10 & 4 \\ 22 & 8 \end{pmatrix}.$$

Thus $\mathbf{A}(2\mathbf{B}) = 2(\mathbf{AB})$, as required.

Solution 2.7

We obtain

$$\mathbf{AB} = \begin{pmatrix} 1 & -1 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 11 \end{pmatrix},$$

$$\mathbf{AC} = \begin{pmatrix} 1 & -1 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix},$$

so

$$\mathbf{AB} + \mathbf{AC} = \begin{pmatrix} -1 \\ 11 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 12 \end{pmatrix}.$$

Also,

$$\mathbf{B} + \mathbf{C} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix},$$

so

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \begin{pmatrix} 1 & -1 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 12 \end{pmatrix}.$$

Thus $\mathbf{AB} + \mathbf{AC} = \mathbf{A}(\mathbf{B} + \mathbf{C})$, as required.

Solution 2.8

For the matrix

$$\mathbf{B} = \begin{pmatrix} 7 & 2.5 & 0 \\ -4 & 1 & 9 \end{pmatrix},$$

(a) $b_{11} = 7$,

(b) $b_{21} = -4$,

(c) $b_{13} = 0$.

Solution 3.1

(a) The birth rates for juveniles and elderly are zero (to 4 decimal places), so we have to consider only the number of workers' babies.

Rearranging equation (3.1) gives, in millions,

$$\begin{aligned} \text{number of births} &= \text{birth rate} \times \text{number of workers} \\ &= 0.0213 \times 37.50 \\ &\simeq 0.7988. \end{aligned}$$

These data lead to an estimate of approximately 800 000 births between June 1990 and June 1991.

(b) To compare the estimated numbers of deaths, we multiply the death rate in each category by the number in the category. This gives, in millions,

$$\begin{aligned} \text{number of worker deaths} &= 0.0031 \times 37.50 \\ &\simeq 0.1163, \end{aligned}$$

$$\begin{aligned} \text{number of elderly deaths} &= 0.0575 \times 8.99 \\ &\simeq 0.5169. \end{aligned}$$

So there were nearly five times as many elderly deaths as those for workers (even though there were about four times as many workers).

Solution 3.2

(a) Table 3.3 tells us that the number that labels the pipe from A_0 to J_1 is the proportionate birth rate for adults. Table 3.2 gives this a value of 0.0172.

(b) Table 3.3 tells us that the number that labels the pipe from A_0 to A_1 corresponds to the proportion of adults who do not die in year 0. Table 3.2 gives the death rate for adults as 0.0136, so the proportion surviving is

$$1 - 0.0136 = 0.9864 \quad \text{(to 4 d.p.)}.$$

(c) Table 3.3 tells us that the number that labels the pipe from J_0 to A_1 corresponds to the proportion of surviving juveniles who reach 15. Table 3.2 gives the death rate for juveniles as 0.0008, so the proportion surviving is $1 - 0.0008 = 0.9992$. Using the assumption that $\frac{1}{15}$ of these juveniles become adults each year gives the label as

$$\frac{1}{15} \times 0.9992 = 0.0666 \quad \text{(to 4 d.p.)}.$$

Solution 3.3

Birth and death rates constant This assumption was needed to create the matrix model in equation (3.2). It is likely to be valid for only short time periods. It is this assumption that limits the period over which the model can predict with any accuracy.

No juveniles give birth This assumption was not necessary to obtain the type of model given in the text. The data in Table 3.2 show that this is a fair assumption to make, since the birth rate for juveniles is insignificant (0.0000 to 4 d.p.).

$\frac{1}{15}$ of surviving juveniles become adults each year This assumption was needed to create the matrix model. It means that the spread of ages within the juvenile subpopulation is uniform. This will be the case if the changes in the population are small compared with the size of the population. A quick check on the predictions of the model will reveal the extent to which this is valid.

No emigration or immigration This factor was completely ignored in the text. More data are needed to decide whether this is a valid assumption. (In fact, this is a fair assumption for the UK population, as the rate of immigration is almost equal to the rate of emigration.)

Solution 3.4

The data in Table 3.2 give the initial values

$$\begin{pmatrix} J_0 \\ A_0 \end{pmatrix} = \begin{pmatrix} 10.92 \\ 46.49 \end{pmatrix}.$$

Now use equation (3.2) to find estimates for year 1:

$$\begin{pmatrix} J_1 \\ A_1 \end{pmatrix} = \begin{pmatrix} 0.9326 & 0.0172 \\ 0.0666 & 0.9864 \end{pmatrix} \begin{pmatrix} 10.92 \\ 46.49 \end{pmatrix} \simeq \begin{pmatrix} 10.98 \\ 46.59 \end{pmatrix}.$$

These estimates for year 1 can be used to calculate estimates for year 2:

$$\begin{pmatrix} J_2 \\ A_2 \end{pmatrix} = \begin{pmatrix} 0.9326 & 0.0172 \\ 0.0666 & 0.9864 \end{pmatrix} \begin{pmatrix} 10.98 \\ 46.59 \end{pmatrix} \simeq \begin{pmatrix} 11.04 \\ 46.68 \end{pmatrix}.$$

These results can be summarised and interpreted as follows. Each of the predictions is expressed in millions of people.

$J_1 \simeq 10.98$: prediction for juveniles in June 1991.
$A_1 \simeq 46.59$: prediction for adults in June 1991.
$J_2 \simeq 11.04$: prediction for juveniles in June 1992.
$A_2 \simeq 46.68$: prediction for adults in June 1992.

(The values of J_2 and A_2 were calculated using the values of J_1 and A_1 produced by the calculator *before* these were rounded. If you used rounded values, then your answers may be slightly different. The more years calculated in this way, the bigger the difference.)

Solution 5.1

The products **AI**, **IA** and **IB** can be formed, with the following results:

$$\mathbf{AI} = \begin{pmatrix} 3 & 4 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -2 & 1 \end{pmatrix} = \mathbf{A},$$

$$\mathbf{IA} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -2 & 1 \end{pmatrix} = \mathbf{A},$$

$$\mathbf{IB} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 7 & -3 \\ 8 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 7 & -3 \\ 8 & 1 & 2 \end{pmatrix} = \mathbf{B}.$$

The product **BI** cannot be formed (because **B** has 3 columns and **I** has 2 rows).

Solution 5.2

(a) According to the assertion, $\mathbf{Ib} = \mathbf{b}$ since the product **Ib** can be formed (**I** has 3 columns and **b** has 3 rows).

(b) Verification:

$$\begin{aligned} \mathbf{Ib} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 \times b_1 + 0 \times b_2 + 0 \times b_3 \\ 0 \times b_1 + 1 \times b_2 + 0 \times b_3 \\ 0 \times b_1 + 0 \times b_2 + 1 \times b_3 \end{pmatrix} \\ &= \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \mathbf{b}. \end{aligned}$$

Solution 5.3

(a) (i) $\mathbf{AB} = \begin{pmatrix} 8 & 5 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & -5 \\ -3 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$

$\mathbf{BA} = \begin{pmatrix} 2 & -5 \\ -3 & 8 \end{pmatrix} \begin{pmatrix} 8 & 5 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$

(ii) $\mathbf{CD} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$

$\mathbf{DC} = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$

(iii) $\mathbf{PQ} = \begin{pmatrix} 3 & -9 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 4 & 9 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix},$

$\mathbf{QP} = \begin{pmatrix} 4 & 9 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & -9 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}.$

(b) Each pair of matrices in parts (a)(i) and (a)(ii), taken in either order, gives the identity matrix as product.

The pair of matrices in part (a)(iii), taken in either order, gives a product which looks rather like the identity matrix. In fact, the product matrix is a scalar multiple of the identity matrix, namely **3I**.

Solution 5.4

Multiplying the two matrices together gives

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & -bc+ad \end{pmatrix}.$$

So the product is $(ad-bc)\mathbf{I}$, the 2×2 identity matrix scalar multiplied by $ad-bc$.

Solution 5.5

(a) $\det \mathbf{A} = 26 - 25 = 1 \neq 0$; hence $\mathbf{A}$ is invertible and

$$\mathbf{A}^{-1} = \begin{pmatrix} 2 & -5 \\ -5 & 13 \end{pmatrix}.$$

(b) $\det \mathbf{B} = 6 - 0 = 6 \neq 0$; hence $\mathbf{B}$ is invertible and

$$\mathbf{B}^{-1} = \tfrac{1}{6}\begin{pmatrix} 2 & 0 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & 0 \\ \frac{5}{6} & \frac{1}{2} \end{pmatrix}.$$

(c) $\det \mathbf{C} = 1 - 1 = 0$; hence $\mathbf{C}$ is non-invertible, that is, $\mathbf{C}^{-1}$ does not exist.

(d) $\det \mathbf{D} = 2.25 - 1.25 = 1 \neq 0$; hence $\mathbf{D}$ is invertible and

$$\mathbf{D}^{-1} = \begin{pmatrix} 1.5 & -2.5 \\ -0.5 & 1.5 \end{pmatrix}.$$

Solution 5.6

(a) To write the equations

$$\begin{aligned} 2x + 3y &= 3, \\ x + 4y &= -1, \end{aligned}$$

in matrix form, first write down the coefficient matrix $\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$. Then write down the matrix form of the equations:

$$\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

(b) The matrix form of the equations

$$\begin{aligned} 2x \quad\quad &= 6, \\ 3x + 6y &= 15, \end{aligned}$$

is

$$\begin{pmatrix} 2 & 0 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \end{pmatrix}.$$

(c) The matrix form of the equations

$$\begin{aligned} \tfrac{3}{5}x - \tfrac{4}{5}y &= 18, \\ \tfrac{4}{5}x + \tfrac{3}{5}y &= -1, \end{aligned}$$

is

$$\begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 18 \\ -1 \end{pmatrix}.$$

Solution 5.7

(a) Activity 5.6(a) gives the matrix form of the equations as

$$\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$2 \times 4 - 3 \times 1 = 5,$$

so the inverse of the coefficient matrix is

$$\tfrac{1}{5}\begin{pmatrix} 4 & -3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{4}{5} & -\frac{3}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix}.$$

Hence

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{4}{5} & -\frac{3}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{4}{5} \times 3 + (-\frac{3}{5}) \times (-1) \\ (-\frac{1}{5}) \times 3 + \frac{2}{5} \times (-1) \end{pmatrix} \\ &= \begin{pmatrix} \frac{15}{5} \\ -\frac{5}{5} \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ -1 \end{pmatrix}. \end{aligned}$$

The solution is $x = 3$, $y = -1$. (You should always check your solution by substituting back into the original equations.)

(b) Activity 5.6(b) gives the matrix form of the equations as

$$\begin{pmatrix} 2 & 0 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$2 \times 6 - 0 \times 3 = 12,$$

so the inverse of the coefficient matrix is

$$\tfrac{1}{12}\begin{pmatrix} 6 & 0 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{4} & \frac{1}{6} \end{pmatrix}.$$

Hence

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{4} & \frac{1}{6} \end{pmatrix}\begin{pmatrix} 6 \\ 15 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} \times 6 + 0 \times 15 \\ (-\frac{1}{4}) \times 6 + \frac{1}{6} \times 15 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 1 \end{pmatrix}. \end{aligned}$$

The solution is $x = 3$, $y = 1$.

(c) Activity 5.6(c) gives the matrix form of the equations as

$$\begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 18 \\ -1 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$\tfrac{3}{5} \times \tfrac{3}{5} - \left(-\tfrac{4}{5}\right) \times \tfrac{4}{5} = 1,$$

so the inverse of the coefficient matrix is

$$\frac{1}{1}\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

Hence

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} 18 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{3}{5} \times 18 + \frac{4}{5} \times (-1) \\ \left(-\frac{4}{5}\right) \times 18 + \frac{3}{5} \times (-1) \end{pmatrix} \\ &= \begin{pmatrix} \frac{50}{5} \\ -\frac{75}{5} \end{pmatrix} \\ &= \begin{pmatrix} 10 \\ -15 \end{pmatrix}. \end{aligned}$$

The solution is $x = 10$, $y = -15$.

Solution 5.8

(a) The matrix form of the equations is

$$\begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$2 \times 6 - (-3) \times (-4) = 0,$$

so the matrix has no inverse. Hence the matrix method does not provide a solution.

Now let us try solving the equations using the elimination method. Start by doubling the first equation in preparation to eliminate x:

$$\begin{aligned} 4x - 6y &= 10, \\ -4x + 6y &= 7. \end{aligned}$$

Now add the two equations to eliminate x; this gives

$$0x - 0y = 17.$$

Obviously, the equation $0 = 17$ has no solution, so the original equations also have no solution.

(b) The matrix form of the equations is

$$\begin{pmatrix} 1 & 2 \\ -3 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -6 \\ 18 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$1 \times (-6) - 2 \times (-3) = 0,$$

so the matrix has no inverse. Hence the matrix method does not provide a solution.

The elimination method can start by tripling the first equation to give

$$\begin{aligned} 3x + 6y &= -18, \\ -3x - 6y &= 18. \end{aligned}$$

Now add the two equations to eliminate x; this gives $0x + 0y = 0$, that is, $0 = 0$. So the elimination method fails to find values for x and y, and this is the same outcome as with the matrix method.

(In fact, the equation $0 = 0$ is true for all values of x and y, and this is an indication that the equations have infinitely many solutions, as pictured in Figure 5.3, following the activity.)

Solution 5.9

(a) The matrix form of the equations is

$$\begin{pmatrix} 5 & -3 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$5 \times 4 - (-3) \times (-1) = 17.$$

This is non-zero, so the coefficient matrix is invertible and the equations have a unique solution.

(b) The matrix form of the equations is

$$\begin{pmatrix} 3 & -1 \\ -9 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -6 \\ 18 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$3 \times 3 - (-1) \times (-9) = 0.$$

This is zero, so the coefficient matrix is non-invertible and the equations do not have a unique solution. (The equations represent the same line.)

(c) The matrix form of the equations is

$$\begin{pmatrix} 6 & -2 \\ 3 & -7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The determinant of the coefficient matrix is

$$6 \times (-7) - (-2) \times 3 = -36.$$

This is non-zero, so the coefficient matrix is invertible and the equations have a unique solution.

Solutions to Exercises

Solution 1.1

(a) (i) $\begin{pmatrix} 0.35 & 0.85 \\ 0.65 & 0.15 \end{pmatrix}\begin{pmatrix} 2 \\ 6 \end{pmatrix}$

$$= \begin{pmatrix} 0.35 \times 2 + 0.85 \times 6 \\ 0.65 \times 2 + 0.15 \times 6 \end{pmatrix}$$
$$= \begin{pmatrix} 0.7 + 5.1 \\ 1.3 + 0.9 \end{pmatrix}$$
$$= \begin{pmatrix} 5.8 \\ 2.2 \end{pmatrix}$$

(ii) $\begin{pmatrix} 0.35 & 0.85 \\ 0.65 & 0.15 \end{pmatrix}\begin{pmatrix} 50 \\ 100 \end{pmatrix}$

$$= \begin{pmatrix} 0.35 \times 50 + 0.85 \times 100 \\ 0.65 \times 50 + 0.15 \times 100 \end{pmatrix}$$
$$= \begin{pmatrix} 17.5 + 85 \\ 32.5 + 15 \end{pmatrix}$$
$$= \begin{pmatrix} 102.5 \\ 47.5 \end{pmatrix}$$

(iii) $\begin{pmatrix} 0.35 & 0.85 \\ 0.65 & 0.15 \end{pmatrix}\begin{pmatrix} 60 \\ 40 \end{pmatrix}$

$$= \begin{pmatrix} 0.35 \times 60 + 0.85 \times 40 \\ 0.65 \times 60 + 0.15 \times 40 \end{pmatrix}$$
$$= \begin{pmatrix} 21 + 34 \\ 39 + 6 \end{pmatrix}$$
$$= \begin{pmatrix} 55 \\ 45 \end{pmatrix}$$

(b) The network is shown in Figure S.5.

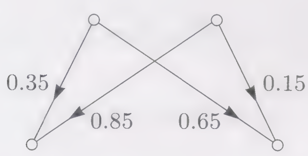


Figure S.5

Solution 2.1

(a) All the products can be formed except **CA** and $\mathbf{D}^2$.

(b) $\mathbf{AB} = \begin{pmatrix} 8 & 12 & 10 \\ 10 & 9 & 14 \end{pmatrix}$, $\mathbf{CD} = \begin{pmatrix} 17 & 8 \\ 14 & 19 \end{pmatrix}$,

$\mathbf{DA} = \begin{pmatrix} 4 & 6 \\ 8 & 0 \\ 3 & 9 \end{pmatrix}$, $\mathbf{DB} = \begin{pmatrix} 8 & 8 & 11 \\ 16 & 24 & 20 \\ 6 & 3 & 9 \end{pmatrix}$,

$\mathbf{A}^2 = \begin{pmatrix} 4 & 0 \\ 5 & 9 \end{pmatrix}$.

Solution 2.2

(a) All the combinations can be formed except $\mathbf{A} + \mathbf{B}$ and $\mathbf{C} + \mathbf{A}$.

(b) $\mathbf{B} + \mathbf{C} = \mathbf{C} + \mathbf{B} = \begin{pmatrix} 5 & 10 & 7 \\ 4 & 4 & 8 \end{pmatrix}$,

$\mathbf{D} + \mathbf{D} = \begin{pmatrix} 2 & 4 \\ 8 & 0 \\ 0 & 6 \end{pmatrix}$,

$\mathbf{A} + \mathbf{A} + \mathbf{A} = \begin{pmatrix} 6 & 0 \\ 3 & 9 \end{pmatrix}$,

$\mathbf{B} - \mathbf{C} = \begin{pmatrix} 3 & 2 & 3 \\ 0 & -2 & -2 \end{pmatrix}$.

Solution 2.3

(a) $5\begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 15 & -5 \\ 0 & 10 \end{pmatrix}$

(b) $\frac{1}{5}\begin{pmatrix} 2 & -5 \\ 10 & 3 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} & -1 \\ 2 & \frac{3}{5} \end{pmatrix}$

(c) $-\frac{2}{3}\begin{pmatrix} 3 & -5 \\ 4 & -9 \end{pmatrix} = \begin{pmatrix} -2 & \frac{10}{3} \\ -\frac{8}{3} & 6 \end{pmatrix}$

Solution 2.4

(a) $\begin{pmatrix} -3 \\ 7 \end{pmatrix} + \begin{pmatrix} 2 \\ -4 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$

(b) $\begin{pmatrix} 0.5 \\ 0.1 \\ 0.4 \end{pmatrix} - \begin{pmatrix} 0.3 \\ 0.4 \\ 0.3 \end{pmatrix} = \begin{pmatrix} 0.2 \\ -0.3 \\ 0.1 \end{pmatrix}$

(c) $\begin{pmatrix} 8 \\ -2 \end{pmatrix} + \begin{pmatrix} -8 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

(d) $\begin{pmatrix} 10 \\ -3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 10 \\ -3 \end{pmatrix}$

Solution 3.1

Let A_n and J_n denote the numbers (in millions) of adults and juveniles, respectively, at the start of year n, where $n = 0$ corresponds to the year starting in January 2000.

(a) The same reasoning as that used for the UK population leads to the following calculations of numbers labelling pipes.

Pipe		
From	To	Label
J_0	J_1	$\frac{14}{15}(1 - 0.0139) \simeq 0.9204$
J_0	A_1	$\frac{1}{15}(1 - 0.0139) \simeq 0.0657$
A_0	J_1	0.0520
A_0	A_1	$1 - 0.0231 \simeq 0.9769$

These labels have been added to the network diagram shown in Figure S.6.

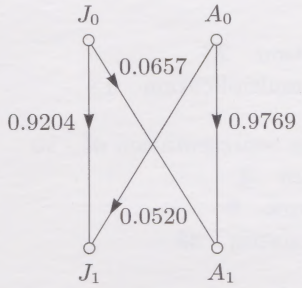


Figure S.6

(b) $\begin{pmatrix} J_{n+1} \\ A_{n+1} \end{pmatrix} = \begin{pmatrix} 0.9204 & 0.0520 \\ 0.0657 & 0.9769 \end{pmatrix} \begin{pmatrix} J_n \\ A_n \end{pmatrix}$

$(n = 0, 1, 2, \dots)$

Solution 5.1

(a) $\det \mathbf{A} = (-3) \times 4 - (-5) \times 1 = -12 + 5 = -7,$

so

$$\mathbf{A}^{-1} = \frac{1}{-7} \begin{pmatrix} 4 & 5 \\ -1 & -3 \end{pmatrix} = \begin{pmatrix} -\frac{4}{7} & -\frac{5}{7} \\ \frac{1}{7} & \frac{3}{7} \end{pmatrix}.$$

(b) $\det \mathbf{B} = 2 \times 5 - 1 \times 5 = 10 - 5 = 5,$

so

$$\mathbf{B}^{-1} = \frac{1}{5} \begin{pmatrix} 5 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{5} \\ -1 & \frac{2}{5} \end{pmatrix}.$$

(c) $\det \mathbf{C} = \frac{3}{5} \times (-\frac{3}{5}) - \frac{4}{5} \times \frac{4}{5}$

$= -\frac{9}{25} - \frac{16}{25} = -\frac{25}{25} = -1,$

so

$$\mathbf{C}^{-1} = \frac{1}{-1} \begin{pmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}.$$

(Note that $\mathbf{C}^{-1} = \mathbf{C}$.)

Solution 5.2

(a) The matrix form of the equations is

$$\begin{pmatrix} -3 & -5 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

From Exercise 5.1(a), the inverse of $\begin{pmatrix} -3 & -5 \\ 1 & 4 \end{pmatrix}$ is $\begin{pmatrix} -\frac{4}{7} & -\frac{5}{7} \\ \frac{1}{7} & \frac{3}{7} \end{pmatrix}$, so

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} -\frac{4}{7} & -\frac{5}{7} \\ \frac{1}{7} & \frac{3}{7} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} (-\frac{4}{7}) \times 1 + (-\frac{5}{7}) \times 2 \\ \frac{1}{7} \times 1 + \frac{3}{7} \times 2 \end{pmatrix} \\ &= \begin{pmatrix} -2 \\ 1 \end{pmatrix}. \end{aligned}$$

The solution is $x = -2$, $y = 1$.

(b) The matrix form of the equations is

$$\begin{pmatrix} 6 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \end{pmatrix}.$$

The inverse of $\mathbf{A} = \begin{pmatrix} 6 & 2 \\ 3 & -1 \end{pmatrix}$ is

$$\mathbf{A}^{-1} = \frac{1}{-12} \begin{pmatrix} -1 & -2 \\ -3 & 6 \end{pmatrix} = \begin{pmatrix} \frac{1}{12} & \frac{1}{6} \\ \frac{1}{4} & -\frac{1}{2} \end{pmatrix},$$

so

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{1}{12} & \frac{1}{6} \\ \frac{1}{4} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 \\ 10 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{12} \times 4 + \frac{1}{6} \times 10 \\ \frac{1}{4} \times 4 + (-\frac{1}{2}) \times 10 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{3} + \frac{5}{3} \\ 1 - 5 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ -4 \end{pmatrix}. \end{aligned}$$

The solution is $x = 2$, $y = -4$.

(c) The matrix form of the equations is

$$\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

The inverse of $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$ is

$$\mathbf{A}^{-1} = \frac{1}{10} \begin{pmatrix} 3 & -2 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{10} & -\frac{1}{5} \\ -\frac{1}{10} & \frac{2}{5} \end{pmatrix},$$

so

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{3}{10} & -\frac{1}{5} \\ -\frac{1}{10} & \frac{2}{5} \end{pmatrix} \begin{pmatrix} 2 \\ -3 \end{pmatrix} \\ &= \begin{pmatrix} \frac{3}{10} \times 2 + (-\frac{1}{5}) \times (-3) \\ (-\frac{1}{10}) \times 2 + \frac{2}{5} \times (-3) \end{pmatrix} \\ &= \begin{pmatrix} \frac{3}{5} + \frac{3}{5} \\ -\frac{1}{5} - \frac{6}{5} \end{pmatrix} \\ &= \begin{pmatrix} \frac{6}{5} \\ -\frac{7}{5} \end{pmatrix}. \end{aligned}$$

The solution is $x = \frac{6}{5}$, $y = -\frac{7}{5}$.

Index

The Open University
ISBN 0 7492 3464 4